EPIGENETICS IN NATURAL ANIMAL POPULATIONS

JUNTAO HU

Department of Biology
McGill University
Montreal, Quebec
July 2018

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Epigenetic mechanisms, especially DNA methylation, have been typically studied in the domains of development and disease; however, recent studies have suggested that epigenetics may also play important roles in ecological and evolutionary processes. In this thesis, I examine the roles of DNA methylation in facilitating animals responding to environmental change. I begin my thesis with a chapter that provides a review of empirical and theoretical studies analysing the effects of epigenetics on phenotypic plasticity and evolution in animals. This forms the background knowledge for the dissertation, and also helps to reveal knowledge gaps to be filled by my other chapters. I found that epigenetic patterns can be shaped by the environment both within and across generations, and that epigenetic variation can play a role in local adaptation. I also evaluate the evolutionary potential of epigenetic variation depending on its autonomy from genetic variation, and its transgenerational stability. During my literature review, I found that environmental effects on epigenetic variation have typically been assessed under laboratory conditions. Thus, to add to the limited but growing body of literature on epigenetics in natural populations of animals, I evaluate epigenetic responses to environmental conditions in three distinct empirical systems and ecological scenarios. In my second chapter, I investigate changes in genome-wide DNA methylation patterns of Trinidadian guppies (*Poecilia reticulata*) during the course of infection by the monogenean ectoparasite, *Gyrodactylus turnbulli*. I found an epigenetic signature of infection by ectoparasites, and identified unique methylation responses at distinct phases of infection. In my third chapter, I explore genome-wide DNA methylation variation of *Anolis* lizards (*Anolis sagrei*) during the early stage of colonization of new habitats. I found that the magnitude of epigenetic variation was dependent on the environmental shift between new and source habitats, and discovered a potential relationship between epigenetic variation and physiological changes in populations at the earliest stages of colonization of new environments. Together with previous work, results from the two chapters suggest that patterns of DNA methylation can rapidly respond to environmental change, and that these methylation changes are involved in the regulation of critical genes and pathways. Although these findings highlight the importance of environmentally-mediated methylation changes, most genomic methylation patterns are static across tissues and throughout life, and some are even stable across generations. Thus, in my last chapter, I use threespine stickleback (*Gasterosteus aculeatus*) to characterise the distribution and function of constitutive

methylation, and assess the amount of heritable methylation. I found a clear pattern of epigenetic variation across generations that is likely to be shaped by genetic variation. In addition, I found that constitutive methylation mapped to genes with functions relevant to fish development, with distinct enrichment of genomic context (promoters or gene bodies) between constitutive hypo- and hypermethylation. Finally, I identified a small but significant amount of heritable methylation. Collectively, my thesis demonstrates the utility of epigenome-wide scans for identifying candidate loci associated with complex phenotypes, and represents a valuable contribution to our understanding of the involvement of epigenetics in ecological and evolutionary process. Consequently, this work helps to improve our ability to predict the capacity of organisms to respond to changing environments.

RÉSUMÉ

Les mécanismes épigénétiques, et plus particulièrement la méthylation de l'ADN, sont généralement étudiés dans les domaines s'intéressant au développement et aux maladies; cependant, de récentes études ont suggéré que l'épigénétique pourrait également jouer un rôle important dans les processus écologiques et évolutifs. Dans cette thèse, j'examine les rôles que joue la méthylation de l'ADN dans la facilitation des réponses environnementales des animaux. Ma thèse commence avec un premier chapitre fournissant une revue des études théoriques et empiriques analysant les effets de l'épigénétique sur la plasticité phénotypique et l'évolution chez les animaux. Cela fournit une bonne connaissance de base pour la dissertation et aide ainsi à révéler les lacunes existantes, qui seront ensuite comblées dans mes chapitres suivants. J'ai mis en évidence que les motifs épigénétiques peuvent être façonnés par l'environnement, que ce soit durant une seule génération ou entre générations, et que la variation épigénétique peut jouer un rôle dans l'adaptation locale. J'ai également évalué le potentiel évolutif de la variation épigénétique en fonction de sa dissociation avec la variation génétique et de sa stabilité transgénérationelle. Cependant, les effets de l'environnement sur la variation épigénétique sont généralement abordés dans des conditions de laboratoire. Ainsi, pour ajouter au sujet grandissant de l'épigénétique dans les populations animales naturelles, j'ai analysé les réponses épigénétiques aux conditions environnementales dans trois systèmes empiriques distincts. Dans mon second chapitre, j'examine les changements au niveau global de l'ADN des schémas de méthylation chez les guppies de Trinidade (*Peocilia reticulata*) au cours d'une infection par l'ectoparatiste monogène, *Gyrodactylus turnbulli*. J'ai révélé l'existence d'une signature épigénétique de l'infection par les ectoparasites et identifié des réponses qui sont uniques à chaque phase distincte de l'infection. Dans mon troisième chapitre, j'explore la variation de la méthylation de l'ADN dans son intégralité chez les lézards *Anolis (Anolis sagrei*) durant les premières phases de colonisation d'un habitat. J'ai trouvé que l'ampleur de la variation épigénétique dépendait du changement environnemental entre les nouveaux habitats et les sources et également découvert une relation potentielle entre la variation épigénétique et les changements physiologiques des populations lors des phases précoces de la colonisation d'un nouvel habitat. Conjointement avec de précédents travaux, les résultats de ces deux premiers chapitres suggèrent que la méthylation de l'ADN peut rapidement répondre aux changements environnementaux et qu'elle est impliquée dans la régulation de gènes et pathways

importants. Bien que l'étude de la variation de la méthylation pouvant être altérée par l'environnement soit intéressante, la plupart des schémas génomiques de méthylation sont statiques entre tissus et au cours de la vie, certains étant même stables à travers les générations. Par conséquent, dans le dernier chapitre de ma thèse, j'utilise l'épinoche (*Gasterosteus aculeatus*) pour caractériser le paysage et la fonction de la methylation constitutive ainsi que pour quantifier l'héritabilité de cette methylation. J'ai trouvé un clair schéma de variation épigénétique inter générationelle vraisemblablement façonnée par la variation génétique. De plus, j'ai démontré que la methylation constitutive est associée à des genes ayant des fonctions importantes pour le développement des poissons, avec un enrichissement distinct du contexte génomique (promoteurs ou corps du gene) entre les hypo- et hypermetylation constitutives. Finalement, j'ai identifié qu'une petite, mais néanmoins importante, quantité de la méthylation est héréditaire. Conjointement, mes travaux de thèse démontrent donc l'utilité de scanner l'entièreté de l'épigénome afin d'identifier des loci candidats associés à de complexes phénotypes et représentent une contribution précieuse à notre compréhension de l'implication de l'épigénétique pour les processus écologiques et évolutifs. Par conséquent, ce travail contribue à améliorer notre capacité à prédire la capacité des organismes à s'adapter aux environnements changeants.

ACKNOWLEDGEMENTS

To start, I would like to thank my great supervisor, Rowan Barrett. I know I was so naïve in science, and I believe my writing was and is still driving him crazy. Rowan always showed his wisdom and patience, and provided the best suggestions and most freedom I ever had and needed during my research, which helped me know how science works, and allowed me to grow as a scientist. I would also like to thank Andrew Hendry for your constant "doubts" on epigenetics, and your tons of "discouragement" before my qualifying exams, which encouraged me to think more on my research. As important as Andrew and Rowan have been to my success over the course of this degree, without the long and decent discussion with Antoine Paccard, I would not start being interested in epigenetics, and without Caroline Leblond we would be without my animal use protocol, money, and a functioning fish lab.

I would like to thank the members of my committee, and those who attended my qualifying exam for their valuable input into my thesis: Melania Cristescu, Daniel Schoen, Neil Price, and Jessica Head. I would also thank many undergraduate students who help with this work either in the field or in the lab: Denice Liu, Areeb Butt, Dominique Dance, Zimeng Li, and Mrinalini Kanyal. Other faculty and students from outside McGill also provided valuable assistance to this research: Sean Rogers, Dave Spiller, Todd Palmer, Sara Smith, and Robert Pringle.

Funding was generously provided by the Natural Sciences and Engineering Research Council of Canada, the Quebec Centre for Biodiversity Science, the Faculty of Science, the McGill Department of Biology, the Redpath Museum (Delise Alison awards), and the China Scholarship Council.

The last four years would have been a lot less enjoyable without the friendship, encouragements, and help of DryBar lab members past and present: Felipe Pérez Jvostov, Kiyoko Gotanda, Krista Oke, Timothy Thurman, Marc-Olivier Beausoleil, Sofia Carvajal, Ananda Martins, Victor Frankel, Jonathas Pereira de Lira, Léa Blondel, Charles Xu, Alan Garcia-Efring (What's up bro!), Arash Askary, Mathilde Salamon, Betzi Peraz, Lotte Skovmand, David Hunt, Daniel Reyes, Grant Haines, Sarah Sanderson, and Chelsea Bishop. I love you guys, and I'll never forget you in the rest of my life!

In the end, I would like to thank my great parents for giving me life, unconditionally supporting my decisions coming to the other half of the earth to pursue my dream. I love you,

dad and mom, and I am really sorry for not being with you for the past four years. Thank you, Wenni, for being my better half, and I could not have gone this far with you.

CONTRIBUTION TO ORIGINAL KNOWLEDGE

All chapters of my thesis constitute original scholarship performed in partial fulfillment of the requirements of the degree of Doctor of Philosophy at McGill University. My first chapter is the first review of empirical studies on epigenetics in natural animal populations, and it also provides an summary of theoretical work that includes epigenetics into evolutionary theories. My second chapter is the first to investigate the effects ectoparasites on host. My third chapter is the first analysis of epigenetic responses in natural animal populations at the early stage of colonization of new habitats. Finally, my last chapter focuses on constitutive and heritable methylation. This chapter is the first to quantify genome-wide methylation patterns across generations in threespine stickleback, characterise the role of constitutive methylation during the fish development, and assess the candidate loci of heritable methylation in fish.

Chapter 1 has been published previously in *Journal of Evolutionary Biology*, Chapter 2 has been published in *Molecular Ecology*. Their common publisher, John Wiley and Sons, has granted permission for it to be reprinted here.

CONTRIBUTIONS OF AUTHORS

I am the first author for all chapters in this thesis. For Chapter 1 and Chapter 2, all co-authors and publishers have granted permission for me to use the manuscripts in this thesis.

Chapter 1 has been published in *Journal of Evolutionary Biology*, and the formatting of the chapter is that of the journal. As stated in the manuscript, I was an integral part of designing the study, collecting the relevant literature, reviewing results from recent studies, and writing the manuscript. My co-authors contributed to all aspects of this chapter. Chapter 2 has been published in *Molecular Ecology*, and the formatting of the chapter is that of the journal. For Chapter 2, I was the primary person behind the design, execution, analysis, and writing. Felipe Pérez-Jvostov and Léa Blondel assisted in experimental set-up and data collection, and both of them contributed to writing the manuscript with input from Rowan Barrett. Chapter 3 and 4 follow the formatting of the journal *Molecular Biology and Evolution*. For Chapter 3, I was the primary person behind the design, bench work, analysis, and writing. Rowan Barrett assisted in developing the idea, analysis, and writing. Arash Askary, Timothy Thurman, Dave Spiller, Todd Palmer, and Rob Pringle assisted in field work, data collection, and the writing of the manuscript. For Chapter 4, I was the primary person for the idea, design, data collection, analysis, and writing. Sara Smith, Tegan Barry, Sean Rogers, and Heather Jamniczhy assisted in sample collection and animal husbandary. Sara Smith and Sean Rogers contributed to the writing of the article with input from Rowan Barrett.

GENERAL INTRODUCTION

The phenotypic variation seen in natural populations today is structured by multiple evolutionary forces and their interactions. Ecologically important phenotypes are often underlain by genetically differences in natural animal populations (Linhart & Grant 1996; Mousseau *et al.* 2000; Merilä & Crnokrak 2001; Mitchell-Olds *et al.* 2007), while the interaction between genotype and environment can also impact phenotypic variation (Via & Lande 1985; DeWitt *et al.* 1998). Therefore, genetic variation and environmental change are important factors for ecological research (e.g., Barrett & Hoekstra 2011; Hoban *et al.* 2016; Charlesworth *et al.* 2017). However, in some cases, there is insufficient genetic variation to permit genetically-based phenotypic responses to environmental change in natural populations (e.g., Przybylo *et al.* 2000; Réale *et al.* 2003; MØller & Merilä 2004; Charmantier *et al.* 2008; Verhoeven & Preite 2014), leaving the question of how populations persist under changing environments unclear.

One potentially important but less understood mechanism that is likely to influence the ecology and evolution of organisms is epigenetics. Epigenetics refers to heritable changes in gene function that cannot be explained by changes in gene sequence alone (Youngson & Whitelaw 2008). This occurs mainly through DNA methylation, histone modification, and small RNA regulation (Duncan *et al.* 2014). DNA methylation, which typically involves the addition of a methyl group to cytosine in CpG context in animals, is perhaps the most extensively characterised epigenetic mechanism in eukaryotes (Jones 2012). Previous studies have suggested that multiple environmental factors, including chemical pollutants (e.g., Bollati *et al.* 2007), dietary components (e.g., Waterland & Jirtle 2003; Sinclair *et al.* 2007) and temperature change (e.g., Navarro-Martín *et al.* 2011; Matsumoto *et al.* 2013) can induce DNA methylation change, and thus affect phenotypes. In addition, recent ecological studies have shown that DNA methylation variation among populations is often associated with ecological factors (e.g., parasite loads, Wenzel & Piertney 2014; temperature and salinity, Huang *et al.* 2017), suggesting the involvement of DNA methylation in local adaptation.

Although our knowledge of the relationship between environmental variation and epigenetics is improving, this understanding has been mostly built through study of lab-reared animals and without explicit consideration of ecological context (Hu & Barrett 2017). My research is motivated by a desire to improve understanding of epigenetics in natural populations of animals, and focuses on a few overarching questions: (1) How do epigenetic

mechanisms impact phenotypic plasticity and adaptive evolution in natural populations of animals? (2) What are the specific epigenetic responses to various environmental stressors (e.g., parasite loads, new habitats) in animals? (3) What is the role of constitutive and heritable methylation in the development and evolution in animals? Addressing these questions will provide new insights for future ecological and evolutionary study into how animals respond to environmental change, as well as add new tools for monitoring and ameliorating the effect of climate change on natural populations.

To provide basic knowledge of epigenetics required for my other chapters, I begin by reviewing recent studies on epigenetically encoded thermal plasticity in animals to illustrate environmentally-mediated epigenetic effects within and across generations. This work provides some of the best examples of the relationship between environmental change and epigenetic variation. I then discuss the role of epigenetic effects during adaptation by exploring population epigenetics in natural animal populations. Finally, I evaluate the evolutionary potential of epigenetic variation depending on its autonomy from genetic variation and its transgenerational stability. Results from this chapter, together with other findings on epigenetics in animals and plants (Bossdorf *et al.* 2008; Feil & Fraga 2012; Verhoeven *et al.* 2016; Richards *et al.* 2017), provide the basis for generating predictive models of the capacity of organisms to adapt to changing climates.

To add to the growing body of literature on epigenetics in natural populations of animals, I evaluate epigenetic responses to environmental conditions in three distinct empirical systems. In my second chapter, I use a well-studied host-parasite system, the guppy *Poecilia reticulata* and its ectoparasitic monogenean *Gyrodactylus turnbulli* to gain mechanistic insight into the dynamics of DNA methylation in host-parasite interactions. Parasites can modulate gene expression profiles in their hosts through epigenetic modifications (Paschos & Allday 2010; Sessions *et al.* 2013), and these modifications can also be associated with an adaptive immune response of the host (Youngblood *et al.* 2010; Boyko & Kovalchuk 2011; Conrath 2011; Holeski *et al.* 2012). However, to date, the most compelling evidence for epigenetic responses to parasite infection has come from studies of endoparasites, such as bacterial pathogens infecting plants (e.g., Dowen *et al.* 2012) and intracellular protozoans infecting vertebrates (e.g., Hari Dass & Vyas 2014). To our knowledge, no study has investigated epigenetic responses of hosts to ectoparasites, or if these responses change during the course of an infection. Ectoparasites are distinct from endoparasites in that they cannot manipulate host cell machinery, and thus cannot directly modify intracellular signalling pathways and host transcription regulation (Cheeseman &

Weitzman 2015). My research for this chapter adds to the large body of literature on guppy-*Gyrodactylus* interactions by characterising the epigenetic modifications associated with infection stage, and demonstrates that epigenetic modifications in guppies play an important role in regulating immune response to *Gyrodactylus*.

In my third chapter, I investigate the role of DNA methylation in facilitating the colonization of new environments, using brown anole (*Anolis sagrei*) individuals transplanted from a source island to either high- or low-quality islands. Previous work has typically focused on the long-term consequences of founder effects and natural selection in colonizing populations, whereas the mechanisms underlying short-term phenotypic responses to colonization of new environments remain relatively unexplored. Furthermore, most studies that have explored the phenotypic consequences of colonization of new habitats have been unable to test if epigenetic mechanisms play a role in the earliest stages of colonization because they have involved comparisons of already established natural populations in source versus colonized environments (e.g., transplanted lizards, Losos *et al.* 1997; Kolbe *et al.* 2012; introduced hose sparrows, Schrey *et al.* 2012; Liebl *et al.* 2013). The results from this chapter suggest that changes in DNA methylation may occur very rapidly after exposure to new environments, and may thus be an important molecular mechanism for mitigating the impact of environmental stressors during colonization of novel habitats.

In my fourth chapter, I examine the roles of constitutive and heritable methylation in the development and evolution of threespine stickleback (*Gasterosteus aculeatus*). While DNA methylation variation can be environmentally responsive (Feil & Fraga 2012), most genomic methylation patterns are static across tissues and throughout life (Smith & Meissner 2013). In addition, although it is clear that DNA methylation patterns can be inherited across generations (Jablonka & Raz 2009; Daxinger & Whitelaw 2012; Lim & Brunet 2013; Heard & Martienssen 2014), current studies have mainly used isogenic lab animals (but see Nätt *et al.* 2012; Weyrich *et al.* 2015). These studies do not typically consider ecological context, and may not reflect epigenetic effects in genetically heterogeneous animals living in the wild. Thus, I assessed the constitutive and heritable methylation patterns in replicated fish populations, and found that epigenetic variation across generations may be shaped by genetic variation. In addition, I showed that constitutive methylation can play important roles in fish development, and that heritable methylation, although only representing a very small proportion of genomic methylation, had a significant genetic basis. These findings demonstrate the importance of constitutive methylation as an epigenetic regulatory

mechanism, and improve our understanding of the heritable basis of population epigenomic variation.

Finally, I conclude my thesis with a brief chapter in which I summarize my general findings and the contributions that my thesis has made to our understanding of how epigenetics works in ecological and evolutionary processes, as well as suggest some directions for future studies.

References

Barrett RD, Hoekstra HE (2011) Molecular spandrels: tests of adaptation at the genetic level. *Nature Reviews Genetics*, **12**, 767-780.

Bollati V, Baccarelli A, Hou L *et al.* (2007) Changes in DNA methylation patterns in subjects exposed to low-dose benzene. *Cancer Research*, **67**, 876-880.

Bossdorf O, Richards CL, Pigliucci M (2008) Epigenetics for ecologists. *Ecology Letters*, **11**, 106-115.

Boyko A, Kovalchuk I (2011) Genetic and epigenetic effects of plant–pathogen interactions: an evolutionary perspective. *Molecular plant*, **4**, 1014-1023.

Carroll SP, Hendry AP, Reznick DN, Fox CW (2007) Evolution on ecological time-scales. *Functional Ecology*, **21**, 387-393.

Charlesworth D, Barton NH, Charlesworth B (2017) The sources of adaptive variation. *Proc Biol Sci*, **284**.

Charmantier A, McCleery RH, Cole LR *et al.* (2008) Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science*, **320**, 800-803.

Conrath U (2011) Molecular aspects of defence priming. *Trends in Plant Science*, **16**, 524-531.

Daxinger L, Whitelaw E (2012) Understanding transgenerational epigenetic inheritance via the gametes in mammals. *Nature Reviews Genetics*, **13**, 153-162.

DeWitt TJ, Sih A, Wilson DS (1998) Costs and limits of phenotypic plasticity. *Trends in Ecology & Evolution*, **13**, 77-81.

Dowen RH, Pelizzola M, Schmitz RJ *et al.* (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.

Duncan EJ, Gluckman PD, Dearden PK (2014) Epigenetics, plasticity, and evolution: How do we link epigenetic change to phenotype? *J Exp Zool B Mol Dev Evol*, **322**, 208-220.

Feil R, Fraga MF (2012) Epigenetics and the environment: emerging patterns and implications. *Nature Reviews Genetics*, **13**, 97-109.

Hairston NG, Ellner SP, Geber MA, Yoshida T, Fox JA (2005) Rapid evolution and the convergence of ecological and evolutionary time. *Ecology Letters*, **8**, 1114-1127.

Hari Dass SA, Vyas A (2014) *Toxoplasma gondii* infection reduces predator aversion in rats through epigenetic modulation in the host medial amygdala. *Molecular Ecology*, **23**, 6114-6122.

Heard E, Martienssen RA (2014) Transgenerational epigenetic inheritance: myths and mechanisms. *Cell*, **157**, 95-109.

Hoban S, Kelley JL, Lotterhos KE *et al.* (2016) Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *American Naturalist*, **188**, 379-397.

Holeski LM, Jander G, Agrawal AA (2012) Transgenerational defense induction and epigenetic inheritance in plants. *Trends in Ecology & Evolution*, **27**, 618-626.

Hu J, Barrett RDH (2017) Epigenetics in natural animal populations. *Journal of Evolutionary Biology*, **30**, 1612-1632.

Huang X, Li S, Ni P *et al.* (2017) Rapid response to changing environments during biological invasions: DNA methylation perspectives. *Molecular Ecology*, **26**, 6621-6633.

Jablonka E, Raz G (2009) Transgenerational Epigenetic Inheritance: Prevalence, Mechanisms, and Implications for the Study of Heredity and Evolution. *The Quarterly Review of Biology*, **84**, 131-176.

Jones PA (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*, **13**, 484-492.

Kaminen-Ahola N, Ahola A, Maga M *et al.* (2010) Maternal ethanol consumption alters the epigenotype and the phenotype of offspring in a mouse model. *PLoS Genetics*, **6**, e1000811.

Kolbe JJ, Leal M, Schoener TW, Spiller DA, Losos JB (2012) Founder effects persist despite adaptive differentiation: a field experiment with lizards. *Science*, **335**, 1086-1089.

Liebl AL, Schrey AW, Richards CL, Martin LB (2013) Patterns of DNA methylation throughout a range expansion of an introduced songbird. *Integrative and Comparative Biology*, **53**, 351-358.

Lim JP, Brunet A (2013) Bridging the transgenerational gap with epigenetic memory. *Trends in Genetics*, **29**, 176-186.

Linhart YB, Grant MC (1996) Evolutionary significance of local genetic differentiation in plants. *Annual Review of Ecology and Systematics*, **27**, 237-277.

Losos JB, Warheit KI, Schoener TW (1997) Adaptive differentiation following experimental island colonization in *Anolis* lizards. *Nature*, **387**, 70.

Matsumoto Y, Buemio A, Chu R, Vafaee M, Crews D (2013) Epigenetic control of gonadal aromatase (*cyp19a1*) in temperature-dependent sex determination of red-eared slider turtles. *PLoS ONE*, **8**, e63599.

Merilä J, Crnokrak P (2001) Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology*, **14**, 892-903.

Mitchell-Olds T, Willis JH, Goldstein DB (2007) Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nature Reviews Genetics*, **8**, 845.

MØller AP, Merilä J (2004) Analysis and interpretation of long-term studies investigating responses to climate change. *Advances in Ecological Research*, **35**, 111-130.

Mousseau TA, Sinervo B, Endler JA (2000) *Adaptive genetic variation in the wild.* Oxford University Press.

Nätt D, Rubin C-J, Wright D *et al.* (2012) Heritable genome-wide variation of gene expression and promoter methylation between wild and domesticated chickens. *BMC Genomics*, **13**, 59.

Navarro-Martín L, Viñas J, Ribas L *et al.* (2011) DNA methylation of the gonadal aromatase (*cyp19a*) promoter is involved in temperature-dependent sex ratio shifts in the European sea bass. *PLoS Genetics*, **7**, e1002447.

Paschos K, Allday MJ (2010) Epigenetic reprogramming of host genes in viral and microbial pathogenesis. *Trends in Microbiology*, **18**, 439-447.

Przybylo R, Sheldon BC, Merilä J (2000) Climatic effects on breeding and morphology: Evidence for phenotypic plasticity. *Journal of Animal Ecology*, **69**, 395-403.

Réale D, McAdam AG, Boutin S, Berteaux D (2003) Genetic and plastic responses of a northern mammal to climate change. *Proceedings of the Royal Society of London B: Biological Sciences*, **270**, 591-596.

Richards CL, Alonso C, Becker C *et al.* (2017) Ecological plant epigenetics: Evidence from model and non-model species, and the way forward. *Ecology Letters*.

Schrey AW, Coon CAC, Grispo MF *et al.* (2012) Epigenetic variation may compensate for decreased genetic variation with introductions: A case study using house sparrows (*Passer domesticus*) on two continents. *Genetics Research International*, **2012**, Article ID 979751.

Sessions OM, Tan Y, Goh KC *et al.* (2013) Host cell transcriptome profile during wild-type and attenuated dengue virus infection. *PLOS Neglected Tropical Diseases*, **7**, e2107.

Sinclair KD, Allegrucci C, Singh R *et al.* (2007) DNA methylation, insulin resistance, and blood pressure in offspring determined by maternal periconceptional B vitamin and

methionine status. *Proceedings of the National Academy of Sciences*, **104**, 19351-19356.

Smith ZD, Meissner A (2013) DNA methylation: roles in mammalian development. *Nature Reviews Genetics*, **14**, 204-220.

Thompson JN (1998) Rapid evolution as an ecological process. *Trends in Ecology & Evolution*, **13**, 329-332.

Verhoeven KJF, Preite V (2014) Epigenetic variation in asexually reproducing organisms. *Evolution*, **68**, 644-655.

Verhoeven KJF, vonHoldt BM, Sork VL (2016) Epigenetics in ecology and evolution: what we know and what we need to know. *Molecular Ecology*, n/a-n/a.

Via S, Lande R (1985) Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution*, **39**, 505-522.

Waterland RA, Jirtle RL (2003) Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Molecular and Cellular Biology*, **23**, 5293-5300.

Wenzel MA, Piertney SB (2014) Fine-scale population epigenetic structure in relation to gastrointestinal parasite load in red grouse (*Lagopus lagopus scotica*). *Molecular Ecology*, **23**, 4256-4273.

Weyrich A, Lenz D, Jeschek M *et al.* (2015) Paternal intergenerational epigenetic response to heat exposure in male Wild guinea pigs. *Molecular Ecology*.

Youngblood B, Davis CW, Ahmed R (2010) Making memories that last a lifetime: heritable functions of self-renewing memory CD8 T cells. *International Immunology*, **22**, 797-803.

Youngson NA, Whitelaw E (2008) Transgenerational epigenetic effects. *Annual Review of Genomics and Human Genetics*, **9**, 233-257.

CHAPTER 1: Epigenetics in natural animal populations

JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.13130

REVIEW

Epigenetics in natural animal populations

J. HU & R. D. H. BARRETT

Redpath Museum and Department of Biology, McGill University, Montreal, QC, Canada

Keywords:

epigenetics;
evolution;
animals;
phenotypic plasticity;
environmental change;
temperature;
stress.

Abstract

Phenotypic plasticity is an important mechanism for populations to buffer themselves from environmental change. While it has long been appreciated that natural populations possess genetic variation in the extent of plasticity, a surge of recent evidence suggests that epigenetic variation could also play an important role in shaping phenotypic responses. Compared with genetic variation, epigenetic variation is more likely to have higher spontaneous rates of mutation and a more sensitive reaction to environmental inputs. In our review, we first provide an overview of recent studies on epigenetically encoded thermal plasticity in animals to illustrate environmentally-mediated epigenetic effects within and across generations. Second, we discuss the role of epigenetic effects during adaptation by exploring population epigenetics in natural animal populations. Finally, we evaluate the evolutionary potential of epigenetic variation depending on its autonomy from genetic variation and its transgenerational stability. Although many of the causal links between epigenetic variation and phenotypic plasticity remain elusive, new data has explored the role of epigenetic variation in facilitating evolution in natural populations. This recent progress in ecological epigenetics will be helpful for generating predictive models of the capacity of organisms to adapt to changing climates.

Abstract

Phenotypic plasticity is an important mechanism for populations to buffer themselves from environmental change. While it has long been appreciated that natural populations possess genetic variation in the extent of plasticity, a surge of recent evidence suggests that epigenetic variation could also play an important role in shaping phenotypic responses. Compared with genetic variation, epigenetic variation is more likely to have higher spontaneous rates of mutation and a more sensitive reaction to environmental inputs. In our review, we first provide an overview of recent studies on epigenetically encoded thermal plasticity in animals to illustrate environmentally-mediated epigenetic effects within and across generations. Second, we discuss the role of epigenetic effects during adaptation by exploring population epigenetics in natural animal populations. Finally, we evaluate the evolutionary potential of epigenetic variation depending on its autonomy from genetic variation and its transgenerational stability. Although many of the causal links between epigenetic variation and phenotypic plasticity remain elusive, new data has explored the role of epigenetic variation in facilitating evolution in natural populations. This recent progress in ecological epigenetics will be helpful for generating predictive models of the capacity of organisms to adapt to changing climates.

Keywords: epigenetics, environmental change, temperature, stress, phenotypic plasticity, evolution, animals

Introduction

Rapid climate change produces a range of new selection pressures on natural populations. As a consequence, depending on the rate and magnitude of environmental change, as well as factors such as habitat fragmentation and natural barriers, many species are experiencing conditions outside their physiological tolerances, and are therefore vulnerable to decline and extinction (Hoffmann & Sgrò, 2011). One important mechanism that may reduce the detrimental effects of environmental change on organisms is phenotypic plasticity; for example, temperature acclimation (Angilletta, 2009) via the adjustment of breeding time in birds (Charmantier *et al.*, 2008) or fibre type composition in the swimming muscles of fish (Scott & Johnston, 2016). While studies on the evolution of phenotypic plasticity have typically used classic quantitative genetics to partition phenotypic variance (V_P) into genetic (V_G), environmental (V_E), and genotype-by-environment variance ($V_{G\times E}$), and focused on how selection acts on genetically-based phenotypic plasticity (Pigliucci, 2005; Chevin *et al.*, 2010; Chevin & Lande, 2010, 2011), it has been suggested that there may be insufficient genetic variation to permit this kind of phenotypic response to climate change in many natural populations (e.g., Przybylo *et al.*, 2000; Réale *et al.*, 2003; Møller & Merilä, 2004; Charmantier *et al.*, 2008). Recently, both empirical and theoretical studies have demonstrated that epigenetic variation can either independently contribute to phenotypic plasticity, or mediate a genetically encoded plastic response (Richards *et al.*, 2010; Duncan *et al.*, 2014). Moreover, several recent findings have shed light on the range of different roles that epigenetic variation may play during evolution. First, unlike genetic variation that is caused by random mutation and is typically independent from environmental change, epigenetic variation may respond to environmental change in some situations (e.g., Waterland & Jirtle, 2003; Kucharski *et al.*, 2008). Second, epigenetic variation may be heritable, though the degree and mechanisms of heritability are not fully understood (e.g., Weaver *et al.,* 2004; Seong *et al.*, 2011). Third, with a higher spontaneous mutation rate than nucleotide mutations (e.g., the epimutation rate was found to be three orders of magnitude higher than the genetic mutation rate in *Arabidopsis thaliana*; Schmitz *et al.*, 2011), depending on its long-term transgenerational stability, epigenetic variation may provide the raw material for phenotypic selection when genetic variation is limited (Becker *et al.*, 2011; Schmitz *et al.*, 2011; Zhang *et al.*, 2013) (Fig. 1.1). These findings suggest that epigenetic variation could play an important role in regulating phenotypic plasticity and facilitating evolutionary adaptation than was previously recognized.

The field of epigenetics has a complex history, beginning in the early 1940s when Waddington first coined the term (Waddington, 1942). As a developmental biologist, Waddington was broadly interested in how genotypes give rise to phenotypes during differentiation and development, with no particular interest in transgenerational events. In recent years, epigenetics has been more narrowly defined to refer to mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in gene sequence (Youngson & Whitelaw, 2008). Epigenetic modifications are mainly based on DNA methylation, histone modification, and small RNAs regulation (Duncan *et al.*, 2014) (Box 1.1). DNA methylation, which typically involves the addition of a methyl group to cytosine within CpG dinucleotides in animals, is perhaps the most extensively characterized epigenetic mechanism in eukaryotes (Jones, 2012). Although DNA methylation has been found in many clades, its pattern and genomic distribution vary widely, suggesting that it may have diverse functions and different modes of targeting specific DNA elements in different taxa (Roberts & Gavery, 2012; Schübeler, 2015). Both histone post-translational modification (PTM) and small RNA regulation can impact gene expression but occur through different mechanisms (Lowdon *et al.*, 2016). Specific histone configurations are known to regulate gene expression by altering the accessibility of the underlying DNA sequences to transcription factors (Zhou *et al.*, 2011) while small RNAs can be partially or fully complementary to mRNAs, resulting in repression or degradation of target sequences (Biggar & Storey, 2015). Although our knowledge of the epigenetic machinery underlying cell signalling (i.e., how cells perceive external and internal signals, and transmit the signals to cellular machinery to activate responses) is rapidly improving, the ecological and evolutionary consequences of different epigenetic mechanisms remain poorly understood.

To achieve a more comprehensive understanding of the roles that epigenetic mechanisms may play in facilitating phenotypic plasticity and evolution, it is important to consider epigenetic processes at the population level (Johnson & Tricker, 2010). A number of recent studies have investigated the epigenetic mechanisms underlying environmentally induced phenotypes under laboratory settings (e.g., in animals: Dolinoy *et al.*, 2007; Kucharski *et al.*, 2008; in plants: Vaugh *et al.*, 2007; Salmon *et al.*, 2008). An important next step will be to study epigenetic processes under field conditions with natural levels of environmental and genetic heterogeneity (Ledon-Rettig, 2013). Here, we first review existing literature on epigenetically encoded plasticity in animals, with a specific focus on thermal plasticity since these studies provide some of the clearest examples for understanding mechanisms of generating population epigenetic variation. Second, we assess the levels of

epigenetic variation in natural animal populations, emphasizing the relationship between epigenetic variation and genetic variation during adaptation. Third, we evaluate the evolutionary potential of epigenetic variation depending on its autonomy from genetic variation and its transgenerational stability. Finally, we review theoretical models that discuss epigenetic inheritance within ecological contexts. We do not cover epigenetic effects on phenotypic plasticity in plants or plant population epigenetics as these topics have already been discussed elsewhere (Richards, 2008; Hirsch *et al.*, 2012; Liu *et al.*, 2015). Furthermore, different epigenetic mechanisms and dynamics of plasticity may exist between plants and animals: sessile plants, unlike typically more mobile animals, cannot move to favourable environments, and plastic responses to biotic and abiotic stresses are more ubiquitous in plants than animals (Agrawal, 2001). Thus, the mechanisms by which epigenetics contributes to plant phenotypic plasticity and adaptation may differ for animals (Suzuki & Bird, 2008; Youngson &Whitelaw, 2008) (Box 1.2). Although there is a rich body of literature on the relationship between epigenetic variation and genetic variation in human populations (e.g., Bell *et al.*, 2012; Gutierrez-Arcelus *et al.*, 2013; Banovich *et al.*, 2014; McRae *et al.*, 2014), we do not discuss epigenetic heritability in humans here as these studies do not typically consider ecological context.

Our discussion is structured around key questions concerning the mechanisms and consequences of epigenetics in evolutionary processes: How does epigenetic variation shape phenotypic plasticity? Is epigenetic variation transgenerationally stable? What is the relationship between epigenetic patterns and adaptation in natural animal populations? Do empirical studies support theoretical models linking epigenetics, phenotypic plasticity, and evolution? And finally, what are the implications of epigenetic variation for the "evolvability" of natural populations in changing environments? Addressing these questions will be useful for identifying gaps in our understanding of epigenetic processes, and provide new scope for future ecological and evolutionary research into how animals may respond to global climate change.

Epigenetically encoded plasticity in animals

Understanding the mechanisms by which animals perceive and respond to environmental signals is of fundamental importance to ecology and evolution. In many recent studies, environmentally induced changes in gene expression have been associated with altered DNA methylation patterns or with altered histone modification (Feil & Fraga, 2012).

In particular, a burgeoning area for research into epigenetic responses to environmental change has been the investigation of epigenetically encoded thermal plasticity in animals. Recent studies have characterized within-generational and potentially transgenerational epigenetic effects, which are two specific mechanisms that generate population variation, and both chromatin- and nucleic acid-based mechanisms have been explored (Table 1.1). Below, we review the key findings from these studies, which provide some of the best examples for understanding the relationships between epigenetic variation, phenotypic plasticity, and evolution in natural populations of animals experiencing changing environmental conditions.

Within-generational epigenetic effects

Studies of within-generational epigenetic effects have shown that epigenetics can regulate diverse phenotypes associated with responses to temperature change. Below we provide a number of examples to help illustrate this diversity. However, it is important to recognize two key points about these studies. First, they typically have not explicitly considered genetic variation when different epigenetic responses were observed, so in cases where multiple populations or families are investigated it is unclear if intraspecific epigenetic variation was induced by different environments or determined by genotype. Second, recent studies have involved lab-reared populations of animals and thus may not accurately reflect the epigenetic processes occurring in more natural settings. The increased genetic and environmental heterogeneity expected in the wild may result in more complex ecological and evolutionary dynamics and outcomes relative to what has been found in the lab.

In some fish and reptile species, sex determination is triggered by temperature changes during gametogenesis (Valenzuela & Lance, 2004). The sex ratio depends on the activity of the gonadal aromatase Cyp19a, a product of the *cyp19a* gene, which irreversibly converts androgens into oestrogens (Navarro-Martin *et al.*, 2011). However, the molecular mechanisms by which temperature during early development influences *cyp19a* expression have remained elusive until recently. Several studies over the last few years have demonstrated that temperature changes can drive epigenetically encoded sex ratio shifts. For example, exposure of European sea bass (*Dicentrarchus labrax*) to high temperature during a critical period in early development led to an increase in DNA methylation at the *cyp19a* promoter region, and resulted in a greater proportion of males (Navarro-Martin *et al.*, 2011). The increased methylation was only found in gonad tissue and not in the brain, and only at the promoter of the *cyp19a* gene and not at the housekeeping gene *β-actin*, suggesting that

sex and temperature differences in methylation levels are both tissue and gene specific. Furthermore, there was no effect of oestrogen treatments on gonadal *cyp19a* promoter methylation level, supporting the relationship between methylation of the promoter and gender bias. Another intriguing result is that several CpGs were found near transcription binding sites at the *cyp19a* promoter, suggesting potential *cis*-regulation on methylation changes. Similar DNA methylation changes have been observed in the red-eared slider turtle (*Trachemys scripta*) (Matsumoto *et al.*, 2013) and the American alligator (*Alligator mississippiensis*) (Parrott *et al.*, 2014), where high temperatures repressed gonadal aromatase expression in embryos or larvae resulting in male-biased populations. However, studies in turtles and alligators have also found different methylation patterns in other sex determination genes. For example, promoter methylation at *SOX9* showed a converse methylation pattern compared to *cyp19a1* in alligator gonads (Parrott *et al.*, 2014). These results suggest that DNA methylation could act as a key mediator integrating temperature into molecular mechanisms that determine sex in some animal species.

Embryonic temperature or temperature at early critical periods during the establishment of thermal control has also been demonstrated to have other long-term phenotypic effects. Campos *et al.* (2013) provided evidence that DNA methylation patterns were associated with a temperature-induced muscle growth change in the Senegalese sole (*Solea senegalensis*). When embryos were reared at low temperatures, there was a significant increase in promoter methylation of a critical myogenesis regulation gene, *myog*, in larvae skeletal muscle due to the action of two DNA methyltransferases, *Dnmt1* and *Dnmt3*. As a result, fish reared at low temperatures produced smaller muscle fibres than fish reared at high temperatures. However, it remains unclear whether the changes in DNA methylation variation were specific to *myog* alone because methylation changes at other loci were not investigated. Johnston *et al.* (2009) provided evidence that in addition to DNA methylation, microRNA expression at different embryonic temperatures can also be associated with the transition from hyperplastic to hypertrophic muscle growth phenotypes in adult zebrafish (*Danio rerio*). Effects of microRNA on thermal plasticity are not only confined to simple developmental transitions, but have also been shown to be involved in complex neuronal network remodelling. For example, demethylation of histone H3 at lysine 27 (H3K27) in the promoter of the brain-derived neurotrophic factor (*Bdnf*) can help build thermotolerance acquisition in chicks (Kisliouk & Meiri, 2009). Chicks injected with a microRNA, miR-138, during a critical period in establishment of thermal control exhibit difficulties in controlling body temperature after being exposed to heat stress. It is thought that the disruption of thermoregulation arises

because miR-138 prevents *Bdnf* promoters from gaining methylation (Kisliouk *et al.*, 2011). Furthermore, antisense knockdown of H3K27-specific lysine histone methyltransferase (HMT), which was correlated with the demethylation of H3K27, has been shown to disrupt thermoregulation establishment and inhibit *Bdnf* mRNA expression (Kisliouk & Meiri, 2009). The above examples suggest that instead of isolated epigenetic mechanisms, it is often suites of epigenetic mechanisms that act in concert to influence animal responses to temperature change.

In addition to developmental transitions, physiological activity in animals is also closely related with temperature change. Marsh & Pasqualone (2014) showed that temperature altered metabolic rates of an Antarctic polychaete, *Spiophanes tcheriniai*, and that these changes were associated with methylation gains at specific CpG sites. Interestingly, metabolic rates at high temperatures returned to control levels after a 4-week acclimation period, which suggests that DNA methylation might be responsible for regulatory shifts that differentiate metabolic activities. Other examples of the association between water temperature and epigenetic patterns include polar fish that exhibit higher global methylation levels than tropical and temperate fish (Varriale & Bernardi, 2006), and differences in nucleolar organization between winter and summer acclimated carp (*Cyprinus carpio*) (Alvarez *et al.*, 2006). However, the precise mechanisms by which phenotypic and epigenetic patterns are linked, and whether the changes in methylation are adaptive in these cases remain unclear.

The majority of studies investigating the relationship between DNA methylation and thermal plasticity have used non-model organisms. This may be due to a lack of recognizable *Dnmt*-like genes and limited DNA methylation patterns in several well-studied model systems (Suzuki & Bird, 2008; Roberts & Gavery, 2012). For example, the worm *Caenorhabditis elegans* essentially lacks DNA methylation, and there is no transposable element methylation in the honey-bee, *Apis mellifera* (Simpson *et al.*, 1986; Wang *et al.*, 2006). Thus far, only a few studies have been conducted in the classic model organism for thermal biology, *Drosophila melanogaster,* and have mainly focused on the associations between histone modification, heat tolerance, and lifespan. These studies have yielded contrasting results regarding the effects of histone biotinylation on phenotypic variation. After comparing the biotinylation levels of the same lysine residues (K9BioH3 and K18BioH3) with controls, Camporeale *et al.* (2006) showed that reduced biotinylation in histones caused by knocking down a major catalytic enzyme (holocarboxylase synthetase, HCS) led to decreased lifespan and heat tolerance in treated flies compared to controls within

one generation. In contrast, although Smith *et al.* (2007) found that flies fed on a biotin-deficient diet for 12 generations also exhibited decreased biotinylated histones, their lifespan and resistance to heat stress actually increased relative to control lines. The divergent results in these two studies may imply that lifespan and heat stress resistance is impacted differently by short-term decreased histone biotinylation versus adaptation to histone biotinylation deficiency over multiple generations. A possible explanation is the hypothetical 'transgenerational wash out' epigenetic effect (Burggren, 2015), where the level of epigenetically-caused phenotypic modification, in this case reduced lifespan and heat stress resistance, progressively declines across generations to sub-detectable levels. This decline may result from rapid adaptation caused by switching between epigenetic variants in periodic environments, as indicated by recent models (Furrow & Feldman, 2014; Kuijper & Johnstone 2016; Uller *et al.*, 2015).

In summary, the current literature investigating within-generational epigenetic effects suggests that temperature changes can strongly influence epigenetic patterns and the phenotypes associated with these epigenetic modifications. However, most studies have not explicitly considered the source of epigenetic variation, for example, environmental or genetic variation, and they have typically been conducted under lab conditions. In addition, almost all studies were conducted within one generation, which has precluded testing of transgenerational epigenetic effects. This is of course important because the evolutionary relevance of epigenetic effects rests on whether the responses are heritable (Richards, 2006; Heard & Martienssen, 2014).

Transgenerational epigenetic effects

Although the resetting of some epigenetic marks at one or more points during an organism's life cycle inhibits the inheritance of epigenetic modifications across generations, a growing number of examples of transgenerational epigenetic inheritance in model systems have been reported in recent years (Jablonka & Raz, 2009; Daxinger & Whitelaw, 2012; Lim & Brunet, 2013). Notably, a number of recent examples have considered transgenerational responses to heat exposure. For example, Seong *et al.* (2011) showed that heat shock can induce the repression of the *white* gene as a result of heterochromatic disruption via drosophila activation transcription factor 2 (dATF-2), a transcription factor functioning in heterochromatin nucleation (Jia *et al.*, 2004), and that the effect is maintained over several generations before returning to the normal state. This result implies that although the

epigenome can be altered under environmental stress, the chromatin state retains its capacity to be reset once the environmental cues that initially induced the variation have disappeared. However, evidence that transgenerational inheritance has an epigenetic basis is generally lacking in mammals, especially in non-model mammals. One exception is Weyrich *et al.* (2016), who exposed adult male guinea pigs (*Cavia aperea*) to increased ambient temperature during spermatogenesis, and allowed them to mate with the same female before and after the heat exposure. There were immediate epigenetic responses to increased temperature in these males, and importantly, modified methylation was also detected in the testes of their sons, which suggests that the epimutations can possibly persist to the F2 generation. This is important because heritable epigenetic effects that contribute to a fitness increase in offspring will clearly have evolutionary consequences. Another example comes from the analysis of CpG depletion in several coral species responding to thermal stress (Dimond & Roberts, 2016). The analysis of CpG depletion is based on the hyper-mutability of methylated cytosines, which readily deaminate to thymine residues over evolutionary time (Roberts & Gavery, 2012). This results in a reduction in CpG dinucleotides, which implies that hypermethylated genomic regions are associated with a reduced number of CpGs, whereas genomic regions enriched with CpGs are hypomethylated. Dimond & Roberts showed that historically hypomethylated regions are enriched in differentially expressed genes that are responsive to thermal stress. These results add to a small but growing body of evidence supporting an association between transgenerational hypomethylation and stress-induced responses (Anway *et al.*, 2005; Luna & Ton, 2012; Luna *et al.*, 2012, Olson & Roberts, 2014). These studies are intriguing because they suggest a possible link between DNA methylation and plastic responses to long-term environmental change. However, it is important to note that this work has not been able to establish causal links between epigenetic variation and fitness differences, and thus it is difficult to know if the observed epigenetic variation has been selected for via evolutionary processes or only represents the stochastic transmission of epimutations. In addition, recent studies have not been designed to distinguish between temperature-induced epigenetic variation that is transmitted only across a single generation and variation that can be inherited for several generations regardless of whether the temperature stress in maintained. Discriminating between these scenarios would help to better understand the evolutionary significance of environmentally-induced transgenerational epigenetic variation in animals.

In summary, studies of epigenetically encoded thermal plasticity have started to tackle several important questions in ecological epigenetics, for example, how does epigenetic

variation shape ecological relevant phenotypes, and is environment induced epigenetic variation transgenerationally stable? Although epigenetic analysis of thermal plasticity is still in its infancy (Box 1.4), results from these studies suggest that methylation patterns can be inherited and also play an important role in transgenerational responses to thermal stress. This work will serve as an important reference for future studies to investigate long-term epigenetic responses to the diverse stressors caused by environmental change.

Population epigenetics in the wild

We now know that epigenetic variation can be triggered by exposure to different environmental conditions, and can sometimes be transmitted across generations. Further understanding of epigenetic processes will be aided by empirical assessment of the amount of population variation that results from either within- or transgenerational epigenetic variation. This is important because epigenetic variation can explain some phenotypic variation that cannot be attributed to genetic variation, and could thereby facilitate responses to environmental change (Bossdorf *et al.*, 2008). However, despite abundant speculation about the potential ecological and evolutionary implications of epigenetic variation, most studies have been carried out on lab-raised animals, and thus the importance of epigenetic processes in natural populations remains unclear. Furthermore, epigenetic variation has been typically studied at the individual level, which has made it difficult to discern its implications for population level evolutionary responses. Evidence that epigenetic variation contributes to adaptation should ultimately come from studies in natural populations (Burggren, 2015). Although most of the well-documented cases of epigenetic variation in nature are from plant populations (Kalisz & Purugganan, 2004; Richards, 2006; Hirsch *et al.*, 2012), recent studies in wild animal populations have also suggested links between epigenetic variation, especially DNA methylation, and local adaptation. Below, we review key findings related to epigenetic variation in wild animal populations.

In most wild animal populations examined to date, there has been an excess of DNA methylation variation relative to genetic variation (Massicotte *et al.*, 2011; Morán & Pérez-Figueroa, 2011; Liu *et al.*, 2012; Massicotte & Angers, 2012; Schrey *et al.*, 2012; Liebl *et al.*, 2013; Skinner *et al.*, 2014; Wenzel & Piertney, 2014). First investigations into epigenetic variation in wild animal populations involved the salmonid, *Oncorhynchus mykiss* (Blouin *et al.*, 2010). The authors tested if distinct levels of DNA methylation variation could explain differential survival rates between fish in two different habitats, but found no significant

differences, possibly due to small sample size (six fish in total) and low resolution methods (MS-AFLP). However, recent studies in clonal fish (*Chrosomus eos-neogaeus*) (Massicotte *et al.*, 2011; Massicotte & Angers, 2012), round-leaf bats (*Hipposideros armiger*) (Liu *et al.*, 2012), house sparrows (*Passer domesticus*) (Schrey *et al.*, 2012; Liebl *et al.*, 2013), Atlantic salmon (*Salmo salar*) (Morán & Pérez-Figueroa, 2011), Darwin's finches (Skinner *et al.*, 2014), red grouse (*Lagopus lagopus scotica*) (Wenzel & Piertney, 2014), *Daphnia* (Schield *et al.*, 2015), yellow baboon (*Papio cynocephalus*) (Lea *et al.*, 2016), Tessellated darter (*Etheostoma olmstedi*) (Smith *et al.*, 2016), and threespine stickleback (*Gasterosteus aculeatus*) (Smith *et al.*, 2015; Trucchi *et al.*, 2016) have all showed population, habitat or species-specific DNA methylation patterns. These patterns may indicate that epigenetic variation is both environmentally sensitive and common among wild animal populations, and could play an important role in regulating phenotypic traits during local adaptation. However, none of the empirical work to date has been designed to assess the degree of autonomy between epigenetic variation and genetic variation. This is important because, as we will discuss in the next section, the effects of epigenetic variation on phenotypic plasticity and evolution can be subsumed into the effects of genetic variation if epimutation is guided by underlying genetic variation. Moreover, most population epigenetic work to date has focused on DNA methylation variation, and it is important to note that other epigenetic mechanisms, for example, histone modification, chromatin remodelling, and noncoding RNAs, may also play important roles in shaping phenotypic variation between populations.

In summary, empirical studies with wild animal populations have demonstrated that epigenetic variation can be documented outside of the laboratory. However, the number of examples is still small, and the traits that natural epigenetic variation has been associated with are largely limited to developmental and morphological phenotypes. Moreover, the extent to which genetic variation controls epigenetic variation, and the stability of population epigenetic variation remain unclear. Further studies that broaden the search for epigenetic variation in natural populations, and assess the importance of such variation for adaptation would be valuable.

The evolutionary potential of epigenetic variation

Experimental studies investigating the role of epigenetic variation in adaptive evolution are in their initial stages (Verhoeven *et al.*, 2016). The evolutionary relevance of epigenetic variation rests on whether epigenetically induced responses are under genetic control

(Richards 2006), and whether epigenetic variation can improve species persistence. Although it is clear that epigenetically induced responses can be inherited over several generations in the laboratory (Jablonka & Raz, 2009; Daxinger & Whitelaw, 2012; Lim & Brunet, 2013), the stability of these responses over longer evolutionary time-scales is unclear. In this section, we will discuss the potential evolutionary impact of epigenetic variation by focusing on two key questions: 1) How autonomous is epigenetic variation from genetic variation? 2) How stable is transgenerational epigenetic variation?

How autonomous is epigenetic variation from genetic variation?

The extent to which epigenetic variation is under genetic control is an important first step in assessing the evolutionary potential of epigenetic processes (Richards, 2006; Bossdorf *et al.*, 2008). To simplify the relationship between genetic and epigenetic variation, Richards (2006) defined three classes of epigenetic variation: obligatory, which is completely dependent on genetic variation (e.g., differentially methylated sites were frequently found within repetitive DNA in dogs; Janowitz Koch *et al.*, 2016); facilitated, which is directed or loosely potentiated by genotype (e.g., *Agouti* viable yellow epialleles in mice; Morgan *et al.*, 1999); and pure, which is typically generated by stochastic events, and is largely independent of genetic variation (e.g., growing divergence in epigenotype during ageing in monozygotic twins; Fraga *et al.*, 2005). Because undetected genetic changes might influence epigenetic variation, it can be difficult to distinguish between pure and facilitated epigenetic variation (Richards, 2006), and thus we mainly discuss the two extremes of epigenetic variation: obligatory and pure (Fig. 1.2).

Both obligatory and pure epigenetic variation could play crucial roles in phenotypic plasticity and evolution, but to date empirical examples of epigenetic variation can largely be categorized as obligatory. Examples include the marginal effects of methylation variation on expression variation when taking SNP effects into account in *Arabidopsis* (Meng *et al.*, 2016), and the targeted methylation of a transposon within the gene *Axin* that induces a unique transcript in a strain of inbred mice (Rakyan *et al.*, 2003). These examples suggest a key role of obligatory epigenetic variation in regulating the active status of transposable elements (TEs), which can be highly sensitive to environmental change, and thus facilitate responses to changing conditions. For instance, elevated temperature alters the expression of piRNAs in *Drosophila melanogaster*. As a result, the mobilization activity of transposons also changes (Brennecke *et al.*, 2008). This process can generate genetic variation and phenotypic

plasticity because high mobility of TEs can increase transposon insertion polymorphisms, and the insertion of TEs into a coding or promoter region can affect gene expression (Rey *et al.*, 2016). Pure epigenetic variation may also help populations respond to environmental change. A recent theoretical study has suggested that when selection acts on pure epigenetic variation as opposed to obligatory epigenetic variation, adaptive phenotypes can occur before genotypic change due to the higher rate of epimutation permitting faster exploration of the fitness landscape (Klironomos *et al.*, 2013). However, none of the empirical work to date has addressed pure epigenetic variation in animals, possibly due to the difficulty of establishing genetically identical populations. Even in plants, there is currently very little data beyond model systems to shed light on the dynamics of pure epigenetic variation during environmental change. A recent study in *Arabidopsis thaliana* suggested that the independent contributions of pure epigenetic variation may be limited, as a large proportion of DNA methylation variants are associated with specific genetic variants (Dubin *et al.*, 2015). Thus, characterizing pure epigenetic variation should be a goal for future work, as it is crucial for understanding whether epigenetic variation can autonomously impact phenotypic variation.

One way to identify epigenetic responses under natural conditions is to look for correlations between epigenetic variation and phenotypic variation or environmental factors that are statistically independent from genetic relatedness of the individuals or populations. Many studies have quantified epigenetic variation by applying standard statistical measures used in population genetics (Box 1.3). To provide conclusive evidence that epigenetic variation can result in ecologically-relevant phenotypic changes that are autonomous from genotypic variation, we suggest that future studies of pure epigenetic variation could transplant different populations into common environments, and test for the contributions of genetic effects to epigenetic variation by testing for genome- or epigenome-wide associations in sample individuals and their offspring (GWAS and EWAS, respectively), after correcting for confounding by genetic background by using a kinship matrix (Dubin *et al.*, 2015; Orozco *et al.*, 2015; Gugger *et al.*, 2016; Lea *et al.*, 2016). Several recent studies have also developed statistical approaches to partition environmental and genetic effects on epigenetic variation. For example, a linear mixed model or binomial mixed effect model which treats environment as a fixed effect, and the contributions from *cis*- and *trans*- genetic variants as random effects, has been used to successfully delineate phenotypic variation into components that are sensitive to temperature treatments (Amanda *et al.*, 2016). Alternatively, a leading principle coordinates analysis can be used when both GWAS and EWAS are available (Rakyan *et al.*, 2011). While these approaches to identify environmental factors are promising, there are

several caveats when interpreting results from such experimental designs. For example, detecting *cis* association between SNP and epigenetic variation does not necessarily imply genetic regulation but may simply be due to linkage disequilibrium (LD) between epigenetic variation at a locus and its proximal SNPs (Taudt *et al.*, 2016). Moreover, the causality between genetic variation and epigenetic variation may be reversed. An emerging view suggests a reciprocal relationship between TEs and epigenetic variation where epigenetic changes can also induce TE-associated genetic variation (Rey *et al.*, 2016), though many detected *cis* associations have been found in the context of SNP-mediated epigenetic silencing of nearby transposable elements (TEs). This reciprocal relationship has implications for inferring mechanisms underlying temperature-induced response. For instance, elevated temperatures were shown to alter the expression of piRNAs in *Drosophila melanogaster*, and as a result, the mobilization activity of transposons also changed (Brennecke *et al*., 2008). If the high mobility of TEs accelerates mutation rates, and the insertion of TEs into a coding or promoter region affects gene expression, the causal explanation for gene expression changes will be epigenetic variation but not genetic regulation.

Because genetic variation can obscure the role of epigenetic variation, simplified experimental systems in which confounding effects of genetic variation have been reduced to a minimum may be useful for isolating the contributions of epigenetic processes. Bossdorf *et al.* (2008) have proposed three approaches to control for the effects of genetic variation when studying the effects of epigenotype on phenotype by either reducing epigenetic or genetic heterogeneity: first, to use species with known deficiencies in epigenetic mechanisms; second, to use demethylating agents to inhibit activities of DNA methyltransferases (Dnmts), and thereby induce experimental demethylation; and third, to choose a study system with a known lack of genetic variation. As mutants with known deficiencies in epigenetic mechanisms do not yet exist for most animal systems, and the use of *in vivo* demethylating agents can lead to undesired side effects caused by untargeted, genome-wide demethylation (Verhoeven *et al.*, 2016), many population epigenetic studies to date have used the last strategy. For example, researchers have used populations with limited genetic variation resulting from clonal reproduction (e.g., clonal fish, Massicotte *et al.*, 2011; Massicotte & Angers, 2012) or bottlenecks following invasion (e.g., house sparrows, Schrey *et al.*, 2012; Liebl *et al.*, 2013). Following from classic investigations in plants (Cubas *et al.*, 1999; Kalisz & Purugganan, 2004), these studies provide the first indications from animals that DNA methylation may sometime act independently from underlying genetic variation, and facilitate a clearer evaluation of the consequences of epigenetic variation. However, mutation

accumulation is still possible in clonal lines, and even the reduced genetic variation in bottlenecked populations may still be sufficient to contribute to epigenetic responses. In summary, although it remains challenging to explicitly partition the genetic basis of epigenetic variation, the ability of autonomous epigenetic variation to casue phenotypic change is increasingly appreciated, and should be considered as a potential mechanism when adaptive traits cannot be explained by common genetic variants.

How stable is transgenerational epigenetic variation?

A major barrier to transgenerational epigenetic inheritance is germline reprogramming. In contrast to plants (Verhoeven *et al.*, 2016), in animals, especially mammals, extensive erasing of epigenetic modifications occurs both in the germline and in the zygote immediately after fertilization. Thus, it is more difficult to inherit epigenetic marks that are not associated with sequence variants across generations in animals. Indeed, heritable epigenetic variation that is independent from genetic control seems to be more common in plants (Taudt *et al.*, 2016), but emerging evidence has shown that such epigenetic inheritance may also exist in animals (Youngson & Whitelaw, 2008; Jablonka & Raz, 2009; Daxinger & Whitelaw, 2012; Lim & Brunet, 2013; Heard & Martienssen, 2014). We distinguish between three main sources of transgenerational epigenetic variation: genetically directed, environment-directed, and stochastic (Shea *et al.*, 2011; Taudt *et al.*, 2016). Genetically directed epigenetic variation is regulated by *cis*- or *trans*-acting genetic variation. Environment-directed epigenetic variation is the result of exposure to current or past environmental factors. In contrast, stochastic epigenetic variation, such as epigenetic drift or epimutation, is more analogous to random genetic mutation, and may arise in when organisms are exposed to stressful environments. These types of epigenetic variation can all be heritable and may share molecular mechanisms (e.g., DNA methylation), but differ in their implications for evolution (Fig. 1.2). When epigenetic variation is genetically directed, the effects of epigenetic variation on phenotype could be considered as a component of the genetic effects. Thus, here we focus on environment-directed and stochastic epigenetic variation that may still be stably transmitted despite not being controlled by genotype.

Typical cases of environment-directed epigenetic variation are epigenetics-mediated phenotypic plasticity in changing environments, resulting in environment-specific phenotypes (Verhoeven & Preite, 2014). Depending on the stability of such epigenetic variation, induced phenotypes can be transmitted to offspring if the epigenetic marks can resist resetting

between generations, but may not persist in organisms after the environmental cue that initially induced the variation has disappeared. Empirical studies have supported the role of environment-directed epigenetic variation in mediating phenotypic plasticity within a single generation, and across generations as we reviewed in the above section. The evolutionary implications and adaptive benefits of within- and between-generation phenotypic plasticity have been discussed in Herman & Sultan (2011). In general, environment-directed epigenetic changes beyond one generation alter adaptive dynamics due to the partial uncoupling of the phenotype from the underlying genotype (Bonduriansky & Day, 2009). Such epigenetic variation could be adaptive if parents can predict the offspring environment to some extent, and the effects of epigenetic variation increase both parental and offspring fitness with low cost (Herman *et al.*, 2014). In contrast to environment-directed epigenetic variation, which is expected to show the same pattern in different individuals with the same genotype when exposed to the same environment, stochastic epigenetic variation can contribute to heritable variation that is shaped by natural selection (Shea *et al.*, 2011), and thus will be indistinguishable from genetic variation in a standard heritability analysis (Johannes *et al.*, 2008; Helanterä & Uller, 2010; Tal *et al.*, 2010). When organisms are maladapted or in stressful environments, stochastic epigenetic variation has the potential to facilitate short-term adaptation by producing phenotypically diverse offspring. This may be favourable by allowing greater exploration of phenotypic space, thereby increasing the probability of producing a phenotype that is closer to the optimum (Pál, 1998; Pál & Miklos, 1999). For longer-term adaptation, in a constant environment, unless the strength of selection is high (Klironomos *et al.*, 2013), stable transmission of stochastic epigenetic variation for many generations will be required for natural selection to produce adaptations based on epiallelic variation (Slatkin, 2009). Importantly, because stressful environments can trigger enhanced epimutation rates (Rapp & Wendel, 2005; Verhoeven *et al.*, 2010), rates of stochastic epimutation may slow after adaptation to the current environment has been achieved.

In a summary, whether epigenetic changes are environment-directed or stochastic is likely to influence their adaptive value, but both sources of epigenetic variation may maintain an adaptive phenotype long enough for new genetic mutations to arise and stabilize the phenotype (Klironomos *et al.*, 2013). Under fluctuating environmental conditions, appropriate rates of epigenetic stability from both types of variation may contribute to transient adaptation by allowing organisms to respond to environmental variation without changing their genomes (Lachmann & Jablonka, 1996; Rando & Verstrepen, 2007; Salathé *et al.*, 2009; Verhoeven & Preite, 2014). However, the costs associated with each source of

variation are different. Costs of environment-directed epigenetic variation mainly accrue through the resources required to maintain sensing machinery, and there is also a potentially detrimental time delay between sensing environmental change and making a phenotypic response (Rando & Verstrepen, 2007), and stochastic epigenetic variation can be costly because it will produce some maladaptive phenotypes in every generation. Both environment-directed and stochastic epigenetic variation may compensate for evolutionary potential that is otherwise constrained by the inability to generate phenotypic variation through recombination or genetic variation (Verhoeven & Preite, 2014). Furthermore, germline resetting may also affect the evolutionary potential of environment-directed and stochastic epigenetic variation in different ways. There is no obvious conflict between environment-directed epigenetic variation and resetting because when resetting of epigenetic marks happens, paternally mediated alterations to these markers can still occur after the resetting process. For example, the typically prolonged relationship between mother and offspring in mammals may result in transgenerational persistence of an environment-directed effect (Weaver *et al.*, 2004). As for stochastic epigenetic variation, whether germline resetting is piecemeal or global will affect its evolutionary potential. When global resetting occurs, the evolutionary potential of stochastic epigenetic variation will be reduced because stochastic epigenetic variation requires stable transmission across generations to form the basis of long-term adaptation. In contrast, if the resetting is piecemeal or incomplete then epigenetic loci can more consistently transmit the impact of natural selection on allelic variation between generations.

Heritable epigenetic mutations and evolution: theoretical approaches

Empirical studies of epigenetic inheritance induced by genetic and environmental perturbations have been reviewed elsewhere (Youngson & Whitelaw, 2008; Jablonka & Raz, 2009; Daxinger & Whitelaw, 2012; Lim & Brunet, 2013; Heard & Martienssen, 2014). Here, we mainly survey theoretical work of epigenetic inheritance within ecological contexts. In general, current theoretical studies have applied two approaches to investigate the effects of epigenetic variation on evolution. In the first approach, the main aim is to investigate the effects of stable levels of epigenetic mutation on the maintenance of genetic or phenotypic variation (Pál, 1998; Pál & Miklos, 1999; Day & Bonduriansky, 2011; Carja & Feldman, 2012; Geoghegan & Spencer 2012, 2013; Klironomos *et al.*, 2013; Kronholm & Collins, 2015; Day, 2016). In the second approach, the switching rate of epigenetic variation between

generations has been identified as an evolutionary variable, which can evolve in response to different environments without interacting with genotypes (Jablonka *et al.*, 1995; Lachmann & Jablonka, 1996; Salathé *et al.*, 2009; Feinberg & Irizarry, 2010; Furrow & Feldman, 2014; Uller *et al.*, 2015; Kuijper & Johnstone, 2016; Table 1.2). In general, models of epigenetic switching rates have concluded that the rate of temporal environmental change is a key factor controlling epigenetic variation. In predictable environments, epigenetic switching rate evolves to approximately the inverse of the length of time between environmental changes (Lachmann & Jablonka, 1996; Salathé *et al.*, 2009). In contrast, under unpredictable environmental conditions, epigenetic variation allows the production of phenotypically diverse offspring, which increases the probability of producing a phenotype that is closer to the optimum, and can make epigenetic switching analogous to a genetically encoded bet-hedging strategy in fluctuating environments (Veening *et al.*, 2008; Day, 2016). This is intriguing because some recent findings studied under the context of bet-hedging may be directly translatable to epigenetic switching (e.g., Kussell & Leibler, 2005; Carja *et al.,* 2014).

Results from models that analyse the interactions between heritable epigenetic variation, genetic variation, and phenotypic variation have suggested that adaptation to changing environments can be accelerated by epigenetic variation in a manner analogous to that proposed for within-generational phenotypic plasticity, which facilitates persistence, followed by genetic assimilation, and a reduction in phenotypic plasticity (Via & Lande, 1985; West-Eberhard, 2003; DeWitt & Scheiner, 2004; Richards, 2006; Lande, 2009). Such adaptations can be enhanced by heritable epigenetic variation, thus helping organisms inhabit novel environments (Jablonka & Raz, 2009). However, the conditions under which stochastic versus environment-directed epigenetic variation may be favoured are unclear because current models have only considered each mechanism in isolation of the other. For example, Furrow & Feldman (2014) suggested that epigenetic variation that is environmentally responsive is advantageous under fluctuating environments because the cost of such epigenetic variation is minimal with stably induced transgenerational phenotypic plasticity. In this case, the authors assumed that the potential cost of maladaptive epigenetic variation is high, and suggested that only epigenetic variation that is environmentally-directed towards the optimal fitness state can be favoured. In contrast, Feinberg & Irizarry (2010) considered a model in which stochastic epigenetic variation was the only source of epigenetic variation, and concluded that it could be an important driver of evolutionary adaptation by increasing the range of phenotypes that could be produced by a given genotype in changing environments.

In summary, although recent studies have made progress in exploring the effects of epigenetic variation during adaptive evolution, initial results have yielded inconclusive messages about the predicted effects of epigenetic variation under environmental change, and the relative importance of environment-directed and stochastic epigenetic variation during adaptation. Further theoretical work is warranted to better understand these issues.

Conclusions

Here we use studies of epigenetically encoded thermal plasticity in animals to provide specific examples for understanding the relationship between epigenetic variation and phenotypic plasticity. We then reviewed the patterns and potential evolutionary consequences of epigenetic variation in wild populations. Specific epigenetic patterns are well documented in some animal populations, but their prevalence and relationships with fitness remain under debate. Moreover, while studies in plants and humans have shown a strong correlation between patterns of epigenetic variation and underlying genetic variants, comparable investigations in wild animals have not yet systematically explored the relative contributions of genetic versus epigenetic variation in explaining the heritability of phenotypic traits. Building upon the characterization of molecular mechanisms underlying epigenetic modifications, a number of recent theoretical studies investigating the stability of heritable epigenetic variation have suggested that transgenerational epigenetic markers can play an important role in increasing the "evolvability" of natural populations in changing environments.

Consideration of epigenetic variation allows an expansion of current concepts of variation and evolution in natural populations to consider additional, non-genetic sources of heritable variation that natural selection may act on. However, even with the progress that we describe here, several challenges remain (Box 1.4). One of the most important applications of increasing knowledge of epigenetic variation may be to address the challenge that only a small proportion of variance in complex traits is explained by common genetic variants (Danchin *et al.*, 2011). By helping to fill the missing heritability gaps between genotypes and phenotypes, epigenetics may aid in predicting evolutionary responses to environmental change. Although epigenetic research in natural animal populations is at an early stage, current studies have built a solid foundation for future work to investigate the role of epigenetic variation in regulating phenotypic plasticity in natural environments, and to link this variation with fitness and long-term evolutionary consequences.

Box 1.1 Glossary

Epigenetic modifications: chromatin and DNA modifications that influence genome function but do not change the underlying DNA sequence.

Epimutation: Heritable stochastic change in chromatin state at a given position or region. In the context of cytosine methylation, epimutations are defined as heritable stochastic changes in the methylation status of a single cytosine or of a region or cluster of cytosines. Such changes do not necessarily imply changes in gene expression.

DNA methylation: the addition of methyl groups, usually to a cytosine base, as a means of chemical DNA modification.

Histone modification: a covalent post-translation change to a histone residue, including lysine acetylation, methylation and ubiquitylation, serine phosphorylation, arginine methylation, etc., each catalysed by one or more protein-modifying enzymes, many of which also have non-histone substrates.

Histone biotinylation: a covalent binding of biotin to distinct lysine residues in histones, catalysed by holocarboxylase synthetase (HCS) and biotinidase (BTD). Histone biotinylation has been implicated in heterochromatin structures, DNA repair, and mitotic chromosome condensation.

Small RNAs: a group of RNAs including microRNA (miRNA) and small interfering RNA (siRNA) that are typically less than 25 nucleotides and can influence gene expression through targeted degradation of mRNA or induction of methylation at complementary DNA sequences.

CpG dinucleotides: a cytosine followed (5'-3') by a guanine. Cytosines at CpG dinucleotides constitute the principal target of DNA methylation in vertebrates. In invertebrates, cytosine methylation also occurs in other sequence contexts such as CHG (where H is any nucleotide except for C).

CpG islands: GC-rich DNA sequences that have a high density of CpG dinucleotides.

DNA methyltransferases (Dnmts): a family of enzymes that catalyse the transfer of a methyl group to DNA. There are typically three Dnmts in animals: Dnmt 1 - responsible for maintaining DNA methylation patterns; Dnmt 3a and 3b - required for de novo methylation.

Methylation-sensitive amplified fragment-length polymorphism (MS-AFLP): a commonly used technique for screening variation in DNA methylation. It can identify genome-wide methylation patterns by replacing standard AFLP restriction enzymes with methylation-sensitive enzymes.

Reduced Representation Bisulfite Sequencing (RRBS): A procedure for single base resolution methylation analysis using bisulfite DNA sequencing of a subsection of a genome.

Epigenetic stability: the persistence of modifications in gene expression and/or epigenetic markers that influence gene expression across generations.

Phenotypic plasticity: the ability of a genotype to yield different phenotypes in response to environmental changes.

Box 1.2 Differences in DNA methylation between animals and plants

DNA methylation is the most well characterized epigenetic mechanism in plants and animals, but there are some important differences in how and where it occurs. Five of the most significant differences between animals and plants are: 1) the presence of non-CpG methylation in plants that is targeted to transposable elements (TEs) is typically regulated by small interfering RNAs (siRNAs) (Mette *et al.* 2000; Chan *et al.* 2004, 2005). 2) The timing of germline separation from somatic tissues is typically different between animals and plants. For example, in mammals, primordial germ cells (PGCs) are derived from the epiblast and arise in the posterior primitive streak during gastrulation. Thus, there is limited time for epigenetic alterations to be transmitted into germline cells. In contrast, there is no early separation of germline and soma in plants, and the gametes are derived from vegetative tissue nearing completion of development. This may provide plants with a greater opportunity for "soft" inheritance than mammals (Youngson & Whitelaw 2008). 3) The targets of DNA methylation between animals and plants are different. In vertebrates, gene-bodies are typically methylated whereas CpG gene promoter regions called CpG islands are often unmethylated (Suzuki & Bird 2008). In invertebrates, methylation predominantly occurs in exons (Feng *et al.* 2010; Zemach *et al.* 2010). In contrast, methylation in plants typically occurs on repetitive DNA elements and TEs. 4) In general, DNA methylation occurs globally in vertebrates, with ~70-80% of cytosines in CpG dinucleotides being methylated (Bird & Taggart 1980). In contrast, plants are more similar to most invertebrates in that they typically have mosaic DNA methylation patterns characterized by domains of heavily methylated DNA interspersed with domains that are methylation free (Tweedie *et al.* 1997; Suzuki & Bird 2008). 5) The transgenerational stability of DNA methylation between animals and plants are different. In mammals, a global reset of DNA methylation occurs both in the germline and in the zygote immediately after fertilization (Heard & Martienssen 2014). In contrast, in plants, most of DNA methylation in CG and CHG (where H is A, C, T) sequence contexts is stable during meiosis and embryogenesis, but CHH methylation is specifically reduced (Calarco *et al.*,2012).

Box 1.3 How to quantify epigenetic variation and associate it with genetic and environmental variation

Genetic and epigenetic estimates of variation are fundamentally different. Genetic variation refers to diversity in allele frequencies between individuals or populations, whereas epigenetic variation refers to the presence or absence of epigenetic markers (e.g., DNA methylation) without implying changes in the underlying DNA sequence. The magnitude of genetic versus epigenetic variation is thus difficult to compare, but patterns of change in estimates of variation can be used to contrast genetic and epigenetic diversity. Initially, researchers have utilized Methylation-sensitive AFLP (MS-AFLP) to identify genome-wide methylation patterns (Schrey *et al.* 2013; Verhoeven *et al.* 2016). MS-AFLP identifies a multi-locus epigenotype for each individual by substituting methylation-sensitive isochizomeric enzymes *Msp*I and *Hpa*II for *Mse*I in a standard AFLP protocol (Vos *et al.* 1995). The enzymes *Msp*I and *Hpa*II have different sensitivities in recognizing cytosine methylation in the CCGG context (Salmon *et al.* 2008). *Msp*I does not cut when the inner cytosine is methylated, and *Hap*II does not cut when either or both cytosines are fully methylated or hemi-methylated (Roberts *et al.* 2007). Using this technique and multivariate statistical approaches, Foust *et al.* (2016) identified DNA methylation patterns (but not genetic variation) that were correlated with environmental gradients. Herrera *et al.* (2016) proposed a new analysis, epigenetic isolation-by-distance (IBD), to infer environmental effects on natural epigenetic variation by using genetic IBD as a null model. Their results suggest that local environments are major drivers of epigenetic spatial structure in populations.

Several recent studies have also applied next generation sequencing techniques to simultaneously analyse the relationships between genetic variation, epigenetic variation, and the environment (e.g., Schmitz *et al.* 2011; Platt *et al.* 2015; Gugger *et al.* 2016). In these studies, genomic DNA treated with sodium bisulfite was sequenced. Sodium bisulfite causes deamination of unmethylated cytosines and results in the conversion of these unmethylated cytosines to uracil, leaving methylated cytosines unconverted (Laird 2010). Untreated genomic DNA was also sequenced to provide information regarding genetic variation. Using markers generated by bisulfite sequencing, several studies have demonstrated a strong correlation between methylation variation and climate variables (Gugger *et al.* 2016; Keller *et al.* 2016).

Despite differences in the type of variation being estimated, some of the standard statistical measures used in population genetics for describing patterns of genetic variation

should be transferable to epigenetic variation. For example, statistics that describe the frequency and diversity of alleles may be applied to epiallelic diversity, and measures such as F_{ST}, which describes genetic population structure (e.g., Liebl *et al.* 2013), *h*, which describes the haplotype diversity (e.g., Richards *et al.* 2012), Analysis of Molecular Variance (AMOVA), which detects population differentiation utilizing molecular markers (e.g., Herrera & Bazaga, 2010), and Principal coordinate analyses (PCoA) can be equally useful in describing population differentiation at the epigenetic level (e.g., Gao *et al.* 2010; Wenzel & Piertney 2014; Preite *et al* 2015). Correlations between population statistics for genetic versus epigenetic variation can potentially be analysed using a Mantel test (e.g., Cervera *et al.* 2002; Wenzel & Piertney 2014; Foust *et al.* 2016). In summary, genome-wide genetic variation and epigenetic variation can be quantified simultaneously and statistical methods can help elucidate the degree of autonomy between epigenetic variation and genetic variation, and the relative importance of genetic and epigenetic variation in facilitating population divergence and adaptation.

Box 1.4 Outstanding questions about epigenetically encoded thermal plasticity

Numerous questions remain regarding epigenetically encoded thermal plasticity in natural animal populations, their ecological and evolutionary importance, and potential implications for population responses to climate change.

- How taxonomically widespread is epigenetic variation in thermal plasticity, and is it linked to particular geographical or environmental gradients?

Substantial differences in epigenetic mechanisms and patterns can exist between and within taxa experiencing changes in temperature (Feng *et al.* 2010; Zemach *et al.* 2010), for example, due to differences in methylation maintenance machinery (Alonso *et al.* 2015; Willing *et al.* 2015), and different strategies to maintain body temperature between ectotherms and endotherms. Epigenetic variation may also be linked to particular life history or habitat features (Herman *et al.* 2014; Verhoeven & Preite 2014). Thus, studies of epigenetic variation should be conducted in a wide range of taxa, and between populations inhabiting different thermal environments to determine whether patterns of epigenetic variation are conserved across deep phylogenies and various habitats.

- Are there particular features that make a system a good model for studying epigenetically encoded thermal plasticity?

In order to study the epigenetic basic of phenotypic variation, it is useful for a system to exhibit a significant degree of phenotypic plasticity when responding to environmental changes (Dimond & Roberts 2016). Good long-term datasets connecting environmental parameters with changes in phenotype and gene expression can also help to place functional work within a broader environmental context. For example, Barshis *et al.* (2013) provided useful long-term data on acclimation of coral species to climate change. In particular, species that can resist reprogramming during meiosis and embryogenesis, and transmit changes in DNA methylation to offspring, will help facilitate understanding of transgenerational epigenetic processes (Duncan *et al.*, 2014). For example, Schield *et al.* (2015) characterized methylation variation of clonal *Daphnia ambigua* in response to fish predator cues, showing consistent changes in the epigenome in successive generations. Finally, a well-annotated genome will clearly aid in improving inferences about the epigenetic basis of phenotypic plasticity and local adaptation. Alternative approaches for analysing DNA methylation variation are also available, for example, using a reference-free approach that combines an optimized RRBS protocol with a tailored bioinformatics pipeline (Klughammer *et al.* 2015), or mapping reads to *de novo* constructed genomes. However, these approaches usually come

with limitations, for example, detecting fewer covered CpGs than using reference-based analysis due to repetitive elements (Klughammer *et al.* 2015) or inaccurate mapping due to assembly errors in *de novo* assembled genomes (Earl *et al.* 2011).

- From descriptive data to causal and quantitative effects in epigenetically encoded thermal plasticity

Genome-wide epigenetic markers, especially DNA methylation, have provided a useful tool for studying thermal plasticity. However, it remains difficult to predict from these descriptive data which of the markers, features, and profiles are indicative of causal and quantitative effects of epigenetics on thermal plasticity. Furthermore, very few studies currently account for the confounding effects of genetic variation in producing thermal plasticity. Thus, we think it will be important to adopt statistical methods (e.g., linear mixed models, binomial mixed models) that account for genetic contributions to thermal plasticity. In addition, in model systems that already have epigenetic candidate loci for thermal plasticity, new experimental approaches (e.g., site-specific DNA methylation editing with a catalytically inactive variant of the Cas9 nuclease; Liu *et al.* 2016) may provide a means to pinpoint functional epigenetic effects on thermal plasticity.

- How are epigenetic changes induced by temperature maintained via the germline, and how is the duration of maintenance determined?

The extent to which environmentally induced epigenetic variants persist across generations remains controversial among evolutionary biologists, especially in mammals where germline resetting is more extensive than plants. It is now clear that some molecular mechanisms, for example, small interfering RNAs (RNAs), Piwi-interacting RNAs (piRNAs), and miRNAs can facilitate epigenetic inheritance via the germline in model animals (Lim & Brunet 2013), however, the generality of these mechanisms in animal thermal plasticity warrants further empirical study. It is also unclear how epigenetic machinery affects the duration of epigenetic changes when organisms face temperature changes. DeWitt *et al.* (1998) suggested potential costs and limits of phenotypic plasticity, and if thermal plasticity is regulated by epigenetic changes, similar costs and limits will also be associated with epigenetic variation, making the trade-offs required for maintenance of epigenetic variation another interesting area to be explored.

- What is the proportion of the epigenome that is found to be under genetic control, the relative contributions of *cis*- and *trans*-acting genetic factors, their average effect sizes, and their mechanisms of action in animal thermal plasticity?

A deep understanding of the heritable basis of population epigenetic variation in animals has come mainly from studies of the relationship between DNA methylation and phenotypic traits, for example, disease in humans. Results from these studies have suggested a predominant correlation between *cis*-acting genetic variants and epigenomic variation (Taudt *et al.* 2016). If this is also true for epigenetically encoded thermal plasticity, levels of genetic variation may determine the vulnerability of organisms to changing temperatures. However, these studies also suggested stochasticity in allele-specific epigenetic variation. Stochastic epigenetic variation may compensate for the loss of phenotypic plasticity (Verhoeven & Preite 2014). Thus, it is necessary to first distinguish between different sources of epigenetic variation using fine resolution sequencing techniques and statistical models, and then study the effects of each epigenetic modification on animal thermal plasticity. Techniques used in human studies (e.g., Chromatin immunoprecipitation followed by sequencing (ChIP-seq)) are still financially prohibitive for many lab groups. More cost-effective approaches, for example, Reduced Representation Bisulphite Sequencing (RRBS), may be a good alternative for collecting data with single-nucleotide resolution. These techniques will allow for more detailed understanding of methylation variation at specific and functionally characterized loci, and will aid epigenome-wide association studies that link epigenetic variation to ecologically-relevant traits such as thermal tolerance.

- Under which temperature regimes will environment-directed versus stochastic epigenetic variation be favoured?

The relative importance of these two types of heritable epigenetic variation has not yet been resolved. Which type of epigenetic variation will be favoured may be determined by the degree of environmental variability, or the strength of natural selection (Verhoeven &Preite 2014). Thurman & Barrett (2016) found that selection on genetic variants was strongest over relatively short timescales (<200 generations), with the greatest magnitude of selection occurring within a single generation. Thus, we predict that stochastic epigenetic variation may be favoured over short time periods when thermal environments are stressful or unpredictable. This is because stochastic epigenetic variation allows the production of phenotypes that are closer to the optimum, or rapid phenotypic switching in fluctuating environments (Veening *et al.* 2008; Day 2016). In contrast, environment-directed epigenetic variation will be favoured over longer time periods of consistent environmental conditions, when the process of acquiring information about the changing climate is less risky. Efforts should be made to develop integrated theoretical models that include both types of epigenetic variation, and consider the strength and timescale of selection.

References

Agrawal, A.A. 2001. Phenotypic plasticity in the interactions and evolution of species. *Science* **294**: 321-326.

Alonso, C., Perez, R., Bazaga, P. & Herrera, C.M. 2015. Global DNA cytosine methylation as an evolving trait: phylogenetic signal and correlated evolution with genome size in angiosperms. *Front. Genet.*, **6**: 4.

Alvarez, M., Quezada, C., Molina, A., Krauskopf, M., Vera, M.I. & Thiry, M. 2006. Ultrastructural changes of the carp *Cyprinus carpio* hepatocyte nucleolus during seasonal acclimatization. *Biol. Cell* **98**: 457-463.

Amanda, J.L., Vigalys, T.P., Durst, P.A.P. & Tung, J. 2016. Maximizing ecological and evolutionary insight from bisulfite sequencing data sets. *bioRxiv* doi: https://doi.org/10.1101/091488

Angilletta, M.J. 2009. *Thermal adaptation*. Oxford University Press, Oxford.

Anway, M.D., Cupp, A.S., Uzumcu, M., & Skinner, M.K. 2005. Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* **308**: 1466-1469.

Banovich, N.E., Lan, X., McVicker, G., van de Geijn, B., Degner, J.F. & Blischak, J.D. *et al.* 2014. Methylation QTLs are associated with coordinated changes in transcription factor binding, histone modification, and gene expression levels. *PLoS Genet.* **10**: e1004663.

Barshis, D., Ladner, J., Oliver, T., Seneca, F., Traylor-Knowles, N. & Palumbi, S. 2013 Genomic basis for coral resilience to climate change. *Proc. Natl Acad. Sci. USA* **110:** 1387-1392.

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

Bell, T.J., Tsai, P., Yang, T., Pidsley, R., Nisbet, J. & Glass, D. *et al.* 2012. Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS Genet.* **8**: e1002629.

Biggar, K.K. & Storey, K.B. 2015. Insight into post-transcriptional gene regulation: stress-responsive microRNAs and their role in the environmental stress survival of tolerant animals. *J. Exp. Biol.* **218**: 1281-1289.

Bird, A.P. & Taggart, M.H. 1980. Variable patterns of total DNA and rDNA methylation in animals. *Nucleic Acid Res.* **8**: 1485-1497.

Blouin, M.S., Thuilier, V., Cooper, B., Amarasinghe, V., Cluzel, L. & Hitoshi, A. *et al.* 2010. No evidence for large differences in genomic methylation between wild and hatchery steelhead *Oncorhynchus mykiss*. *Can. J. Fish Aquat. Sci.* **67**: 217-224.

Bonduriansky, R. & Day, T. 2009. Nongenetic inheritance and its evolutionary implications. *Annu. Rev. Ecol. Evol. Syst.* **40**: 103-125.

Bossdorf, O, Richards, C.L. & Pigliucci, M. 2008. Epigenetics for ecologist. *Ecol. Lett.* **11**: 106-116.

Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A. & Hannon, G.J. 2008. An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* **322**: 1387-1392.

Burggren, W.W. 2015. Dynamics of epigenetic phenomena: intergenerational and intragenerational phenotype 'washout'. *J. Exp. Biol.* **218**: 80-87.

Calarco, J.P., Borges, F., Donoghue, M.T., Van Ex, F., Jullien, P.E., Lopes, T. *et al.* 2012. Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. *Cell* **151**: 194-205.

Camporeale, G., Giordano, E., Rendina, R., Zempleni, J., Eissenberg, J.C. 2006. *Drosophila melanogaster* holocarboxylase synthetase is a chromosomal protein required for normal histone biotinylation, gene transcription patterns, lifespan, and heat tolerance. *J. Nutr.* **136**: 2735-2742.

Campos, C., Valente, L.M., Conceição, L.E., Engrola, S. & Fernandes, J.M. 2013. Temperature affects methylation of the *myogenin* putative promoter, its expression and muscle cellularity in Senegalese sole larvae. *Epigenetics* **8**: 389-397.

Carja, O & Feldman, M.W. 2012. An equilibrium for phenotypic variance in fluctuating environments owing to epigenetics. *J. R. Soc. Interface* **9**: 612-623.

Carja. O., Liberman, U. & Feldman, M.W. 2014. The evolution of phenotypic switching in subdivide populations. *Genetics* **196**: 1185-1197

Cervera, M.T., Ruiz-García, L. & Martínez-Zapater, J.M. 2002. Analysis of DNA methylation in *Arabidopsis thaliana* based on methylation-sensitive AFLP markers. *Mol. Genet. Genomics* **268**: 543-552.

Chan, S.W., Henderson, I.R. & Jacobsen, S.E. 2005. Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nat. Rev. Genet.* **6**: 351-360.

Chan, S.W., Zilberman, D., Xie, Z., Johansen, L.K., Carrington, J.C. & Jacobsen, S.E. 2004. RNA silencing genes control de novo DNA methylation. *Science* **303**: 1336.

Charmantier, A., McCleery, R.H., Cole, L.R. & Perrins, C. 2008. Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science* **320**: 800-803.

Chevin, L.M. & Lande, R. 2010. When do adaptive plasticity and genetic evolution prevent extinction of a density-regulated populations? *Evolution* **64**: 1143-1150.

Chevin, L.M., Lande, R. & Mace, G.M. 2010. Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLoS Biol.* **8**: e1000357.

Chevin, L.M. & Lande, R. 2011. Adaptation to marginal habitats by evolution of increased phenotypic plasticity. *J. Evolution Biol.* **24**: 1462-1476.

Cubas, P., Vincent, C. & Coen, E. AN epigenetic mutation responsible for natural variation in floral symmetry. *Nature* **401**: 157-161.

Danchin, É., Charmantier, A., Champagne, F.A., Mesoudi, A., Pujol, B. & Blanchet, S. 2011. Beyond DNA: integrating inclusive inheritance into an extended theory of evolution. *Nat. Rev. Genet.* **12**: 475-486.

Daxinger, L. & Whitelaw, E. 2012. Understanding transgenerational epigenetic inheritance via the gametes in mammals. *Nat. Rev. Genet.* **13**: 153-162.

Day, T. 2016. Interpreting phenotypic antibiotic tolerance and persister cells as evolution via epigenetic inheritance. *Mol. Ecol.* **25**: 1869-1882.

Day, T. & Bonduriansky, R. 2011. A unified approach to the evolutionary consequences of genetic and nongenetic inheritance. *Am. Nat.* **178**: E18-E36.

DeWitt, T.J. & Scheiner, S. 2004. *Phenotypic Plasticity: Functional and Conceptual Approaches,* Oxford University Press, Oxford, UK.

DeWitt, T.J., Sih, A. & Wilson, D.S. 1998. Costs and limits of phenotypic plasticity. *Trend Ecol. Evol.* **13**: 77-81.

Dimond, J.L. & Roberts, S.B. 2016. Germline DNA methylation in reef corals: patterns and potential roles in response to environmental change. *Mol. Ecol.* **25**: 1895-1904.

Dolinoy, D.C., Huang, D. & Jirtle, R.L. 2007. Maternal nutrient supplementation counteracts bisphenol A-induced DNA hypomethylation in early development. *Proc. Natl Acad. Sci. USA* **104**: 13056-13061.

Dubin, M.J., Zhang, P., Meng, D., Remigereau, M.S., Osborne, E.J. & Paolo, C.F. *et al.* 2015. DNA methylation in *Arabidopsis* has a genetic basis and shows evidence of local adaptation. *eLife* **4**: e05255.

Duncan, E.J., Gluckman, P.D. & Dearden, P.K. 2014. Epigenetics, plasticity, and evolution: How do we link epigenetic change to phenotype? *J. Exp. Zool. B Mol. Dev. Evol.* **322**: 208-220.

Earl, D., Bradnam, K., John, J.S., Darling, A., Lin, D., Fass, J. *et al.* 2011. Assemblathon 1: a competitive assessment of de novo short read assembly methods. *Genome Res.* **21**: 2224-2241.

Feil, R. & Fraga, M.F. 2012. Epigenetics and the environment: emerging patterns and implications. *Nat. Rev. Genet.* **13**: 97-109.

Feinberg, A.P. & Irizarry, R.A. 2010. Stochastic epigenetic variation as a driving for of development, evolutionary adaptation, and disease. *Proc. Natl. Acad. Sci. USA* **107**: 1757-1764.

Feng, S., Cokus, S.J., Zhang, X., Chen, P.Y., Bostick, M. & Goll, M.G. *et al.* 2010. Conservation and divergence of methylation patterning in plants and animals. *Proc. Natl. Acad. Sci. USA* **107**: 8689-8694.

Foust, C.M., Preite, V., Schrey, A.W., Alvarez, M., Robertson, M.H. & Verhoeven, K.J.F. *et al.* 2016. Genetic and epigenetic differences associated with environmental gradients in replicate populations of two salt marsh perennials. *Mol. Ecol.* **25**: 1639-1652.

Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F. & Ballestar, M.L. *et al.* 2005. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl Acad. Sci. USA* **102**: 10604-10609.

Furrow, R.E. & Feldman, M.W. 2014. Genetic variation and the evolution of epigenetic regulation. *Evolution* **68**: 673-683.

Gao, L., Geng, Y., Li, B., Chen, J. & Yang, J. 2010. Genome-wide DNA methylation alterations of *Alternanthera philoxeroides* in natural and manipulated habitats: implications for epigenetic regulation of rapid responses to environmental fluctuation and phenotypic variation. *Plant Cell Environ.* **33**: 1820-1827.

Geoghegan, J.L. & Spencer, H.G. 2012. Population-epigenetic models of selection. *Theor. Pop. Biol.* **81**: 232-242.

Geoghegan, J.L. & Spencer, H.G. 2013. The adaptive invasion of epialleles in a heterogeneous environment. *Theor. Pop. Biol.* **88**: 1-8.

Gugger, P., Fitz-Gibbon, S., Pellegrini, M. & Sork, V.L. 2016. Species-wide patterns of DNA methylation variation in Quercus lobate and its association with climate gradients. *Mol. Ecol.* **25**: 1665-1680.

Gutierrez-Arcelus, M., Lappalainen, T., Montgomery, S.B., Buil, A., Halit, O. & Yurovsky, A. *et al.* 2013. Passive and active DNA methylation and the interplay with genetic variation in gene regulation. *eLife* **2**: e00523.

Heard, E. & Martienssen, R.A. 2014. Transgenerational epigenetic inheritance: myths and mechanisms. *Cell* **157**: 95-109.

Helanterä, H. & Uller, T. 2010. The Price equation and extended inheritance. *Philos. Theor. Biol.* **2**: e101.

Herman, J.J., Spencer, H.G., Donohue, K. & Sultan, S.E. 2014. How stable 'should' epigenetic modifications be? Insights from adaptive plasticity and bet hedging. *Evolution* **68**: 632-643.

Herman, J.J. & Sultan, S.J. 2011. Adaptive transgenerational plasticity in plants: case studies, mechanisms, and implications for natural populations. *Front. Plant Sci.* **2**: 102.

Herrera, C.M. & Bazaga, P. 2010. Epigenetic differentiation and relationship to adaptive genetic divergence in discrete populations of the violet *Viola cazorlensis*. *New Phytol.* **187**: 867-876.

Herrera, C.M., Medrano, M. & Bazaga, P. 2013. Epigenetic differentiation persists after male gametogenesis in natural populations of the perennial herb *Helleborus foetidus* (Ranunculaceae). *PLoS ONE* **8**: e70730.

Herrera, C.M., Medrano, M. & Bazaga, P. 2016. Comparative spatial genetics and epigenetics of plant populations: heuristic value and a proof of concept. *Mol. Ecol.* **8**: 1653-1664.

Hirsch, S., Baumberger, R. & Grossniklaus, U. 2012. Epigenetic variation, inheritance, and selection in plant populations. *Cold Spring Harb. Sym.* **77**: 97-104.

Ho, M.W., Tucker, C., Keeley, D. & Saunders, P.T. 1983 Effects of successive generations of ether treatment on penetrance and expression of the *Bithorax* phenocopy in *Drosophila melanogaster*. *J. Exp. Zool.* **225**: 357-368.

Hoffmann, A.A. & Sgrò, C.M. 2011. Climate change and evolutionary adaptation. *Nature* **470**: 479-485.

Jablonka, E., Oborny, B., Molnár, I., Kisdi, E, Hofbauer, J. & Czárán, T. 1995. The adaptive advantage of phenotypic memory in changing environments. *Phil. Trans. R. Soc. Lond. B* **350**: 133-141.

Jablonka, E. & Raz, G. 2009. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *Q. Rev. Biol.* **84**: 131-176.

Janowitz Koch, I., Clark, M.M., Thompson, M.J., Deere-Machemer, K.A., Wang, J. & Duarte L. *et al.* 2016. The converted impact of domestication and transposon insertions on methylation patterns between dogs and grey wolves. *Mol. Ecol.* **25**: 1838-1855.

Jia, S., Noma, K. & Grewal, S.I. 2004. RNAi-independent heterochromatin nucleation by the stress-activated ATF/CREB family proteins. *Science* **304**: 1971-1976.

Johannes, F., Colot, V. & Jansen, R.C. 2008. Epigenome dynamics: a quantitative genetics perspective. *Nat. Rev. Genet.* **9**: 883-890.

Johnson, L.J. & Tricker, P.J. 2010. Epigenomic plasticity within populations: its evolutionary significance and potential. *Heredity* **105**: 113-121.

Johnston, I.A., Lee, H.T., Macqueen, D.J., Paranthaman, K., Kawashima, C. & Anwar, A. *et al.* 2009. Embryonic temperature affects muscle fibre recruitment in adult zebrafish: genome-wide changes in gene and microRNA expression associated with the transition from hyperplastic to hypertrophic growth phenotypes. *J. Exp. Biol.* **212**: 1781-1793.

Jones, P.A. 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**: 484-492.

Kalisz, S. & Purugganan, M.D. 2004. Epialleles via DNA methylation: consequences for plant evolution. *Trends Ecol. Evol.* **19**: 309-314.

Keller, T.E., Lasky, J.R. & Yi, S.V. 2016. The multivariate association between genome-wide DNA methylation and climate across the range of *Arabidopsis thaliana*. *Mol. Ecol.* **25**: 1823-1837.

Kisliouk, T. & Meiri, N. 2009. A critical role for dynamic changes in histone H3 methylation at the *Bdnf* promoter during postnatal thermotolerance acquisition. *Eur. J. Neurosci.* **30**: 1909-1922.

Kisliouk, T., Yosefi, S. & Meiri, N. 2011. MiR-138 inhibits EZH2 methyltransferase expression and methylation of histone H3 at lysine 27, and affects thermotolerance acquisition. *Eur. J. Neurosci.* **33**: 224-235.

Klironomos, F., Berg, J. & Collins, S. 2013. How epigenetic mutations can affect genetic evolution: model and mechanism. *BioEssays* **35**: 571-578.

Klughammer, J., Datlinger, P., Printz, D., Sheffield, N.C., Farlik, M. & Hadler, J. *et al.* 2015. Differential DNA methylation analysis without a reference genome. *Cell Rep.* **13**: 2621-2633.

Kronholm, I. & Collins, S. 2015. Epigenetic mutations can both help and hinder adaptive evolution. *Mol. Ecol.* **25**: 1856-1868.

Kucharski, R., Maleszka, J., Foret, S. & Maleszka, R. 2008. Nutritional control of reproductive status in honeybees via DNA methylation. *Science* **319**: 1827-1830.

Kuijper, B. & Johnstone, R.A. 2016. Parental effects and the evolution of phenotypic memory. *J. Evol. Biol.* 29: 265-276.

Kussell, E., Leibler, S. 2005. Phenotypic diversity, population growth, and information in fluctuating environments. *Science* **309**: 2075-2078.

Lachmann, M. & Jablonka, E. 1996. The inheritance of phenotypes: an adaptation to fluctuating environments. *J. Theor. Biol.* **181**: 1-9.

Laird, P.W. 2010. Principles and challenges of genome-wide DNA methylation analysis. *Nat. Rev. Genet.* **11**: 191-203.

Lande, R. 2009. Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *J. Evol. Biol.* **22**: 1425-1446.

Lea, A.J., Altmann, J., Alberts, S.C. & Tung, J. 2016. Resource base influences genome-wide DNA methylation levels in wild baboons *Papio cynocephalus*. *Mol. Ecol.* **25**: 1681-1696.

Ledon-Rettig, C.C. 2013. Ecological epigenetics: an introduction to the symposium. *Integr. Comp. Biol.* **53**: 307-318.

Lewis, K. 2010. Persister cells. *Annu. Rev. Microbiol.* **64**: 357-372.

Liebl, A.L., Schrey, A.W., Richards, C.L. & Martin, L.B. 2013. Patterns of DNA methylation throughout a range expansion of an introduced songbird. *Integr. Comp. Biol.* **53**: 351-358.

Lim, J.P. & Brunet, A. 2013. Bridging the transgenerational gap with epigenetic memory. *Trends Genet.* **29**: 176-186.

Liu, J., Feng, L., Li, J. & He, Z. 2015. Genetic and epigenetic control of plant heat responses. *Front Plant Sci.* **6**: 267.

Liu, S., Sun, K., Jiang, T., Ho, J.P., Liu, B. & Feng, J. 2012. Natural epigenetic variation in the female great roundleaf bat *Hipposideros armiger* populations. *Mol. Genet. Genomics* **287**: 643-650.

Liu, X.S., Wu, H., Ji, X., Stelzer, Y., Wu, X., Czauderna, S. *et al.* 2016. Editing DNA methylation in the mammalian genome. *Cell* **167**: 233-247.

Lowdon, R.F., Jang, H.S. & Wang, T. 2016. Evolution of epigenetic regulation in vertebrate genome. *Trends Genet.* **32**: 269-283.

Luna, E. & Ton, J. 2012. The epigenetic machinery controlling transgenerational systemic acquired resistance. *Plant Signal. Behav.* **7**: 615-618.

Luna, E., Bruce, T.J., Roberts, M.R. & Ton, J. 2012. Next-generation systemic acquired resistance. *Plant Physiol.* **158**: 844-853.

Marsh, A.G. & Pasqualone, A.A. 2014. DNA methylation and temperature stress in an Antarctic polychaete, *Spiophanes tcherniai*. *Front. Physiol.* **5**: 173-181.

Massicotte, R. & Angers, B. 2012. General-purpose genotype or how epigenetics extend the flexibility of a genotype. *Genet. Res. Int.* **2012**: Article ID 317175.

Massicotte, R., Whitelaw, E. & Angers, B. 2011. DNA methylation: A source of random variation in natural populations. *Epigenetics* **6**: 421-427.

Matsumoto, Y., Buemio, A., Chu, R., Vafaee, M. & Crews, D. 2013. Epigenetic control of gonadal aromatase *cpy19a1* in temperature-dependent sex determination of read-eared slider turtles. *PLoS ONE* **8**: e63599.

McRae, A.F., Powell, J.E., Henders, A.K., Bowdler, L., Hemani, G. & Shah, S. *et al.* 2014. Contribution of genetic variation to transgenerational inheritance of DNA methylation. *Genome Biol.* **15**: R73.

Meng, D., Dubin, M., Zhang, P., Osborne, E.J., Stegle, O. & Clark, R.M. *et al.* 2016. Limited contribution of DNA methylation variation to expression regulation in *Arabidopsis thaliana*. *PLoS Genet.* **12**: e1006141.

Mette, M.F., Aufsatz, W., van der Winden, J, Matzke, M.A. & Matzke, A.J. 2000. Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *EMBO J.* **19**: 5194-5201.

Morgan, H.D., Sutherland, H.G., Martin, D.I. & Whitelaw, E. 1999. Epigenetic inheritance at the agouti locus in the mouse. *Nat. Genet.* **23**: 314-318.

Møller, A.P. & Merilä, J. 2004. Analysis and interpretation of long-term studies investigating responses to climate change. *Adv. Ecol. Res.* **35**: 111-130.

Morán, P. & Pérez-Figueroa, A. 2011. Methylation changes associated with early maturation stages in the Atlantic salmon. *BMC Genet.* **12**: 86.

Navarro-Martin, L., Viñas, J., Ribas, L., Díaz, N., Gutiérrez, A. & Di Croce, L. *et al.* 2011. DNA methylation of the gonadal aromatase *cpy19a* promoter is involved in temperature-dependent sex ratio shifts in the European sea bass. *PLoS Genet.* **7**: e1002447.

Olson, C.E. & Roberts, S.B. 2014. Indication of family-specific DNA methylation patterns in developing oysters. *bioRxiv* doi: http://doi.org/10.1101/012831

Orozco, L.D., Morselli, M., Rubbi, L., Guo, W., Go, J. & Shi, H. *et al.* 2015. Epigenome-wide association of liver methylation patterns and complex metabolic traits in mice. *Cell Metab.* **21**: 905-917.

Pál, C. 1998. Plasticity, memory and the adaptive landscape of the genotype. *Proc. R. Soc. Lond. B* **265**: 1319-1323.

Pál, C. & Miklos, I. 1999 Epigenetic inheritance, genetic assimilation and speciation. *J. Theor. Biol.* **200**: 19-37.

Parrott, B.B., Kohno, S., Cloy-McCoy, J.A. & Guillette, L.J. Jr. 2014. Differential incubation temperatures result in dimorphic DNA methylation patterning of the *SOX9* and aromatase promoters in gonads of alligator *Alligator mississippiensis* embryos. *Biol. Reprod.* **90**: 1-11.

Pigliucci, M. 2005. Evolution of phenotypic plasticity: where are we going now? *Trends Ecol. Evo.* **20**: 481-486.

Platt, A., Gugger, P.F., Pellegrini, M., & Sork, V.L. 2015. Genome-wide signature of local adaptation linked to variable CpG methylation in oak populations. *Mol. Ecol.* **24**: 3823-3830.

Preite, V., Snoek, L.B., Oplaat, C., Biere, A., van der Putten, W.H. & Verhoeven, K.J. 2015. The epigenetic footprint of poleward range-expanding plants in apomictic dandelions. *Mol. Ecol.* **24**: 4406-4418.

Przybylo, R., Sheldon, B.C. & Merilä, J. 2000. Climatic effects on breeding and morphology: evidence for phenotypic plasticity. *J. Anim. Ecol.* **69**: 395-403.

Rakyan, V.K., Chong, S., Champ, M.E., Cuthbert, P.C., Mrogan, H.D. & Luu, K.V. *et al.* 2003. Transgenerational inheritance of epigenetic states at the murine $Axin^{Fu}$ allele occurs after maternal and paternal transmission. *Proc. Natl. Acad. Sci. USA* **100**: 2538-2543.

Rakyan, V.K., Down, T.A., Balding, D.J. & Beck S. 2011. Epigenome-wide association studies for common human diseases. *Nat. Rev. Genet.* **12:** 529-541.

Rando, O.J. & Verstrepen, K.J. 2007. Timescales of genetic and epigenetic inheritance. *Cell* **128**: 655-668.

Rapp, O.J. & Wendel, J.F. 2005. Epigenetics and plant evolution. *New Phytol.* **168**: 81-91.

Réale, D., McAdam, A.G., Boutin, S. & Berteaux, D. 2003. Genetic and plastic responses of a northern mammal to climate change. *Proc. R. Soc. B* **270**: 591-596.

Rey, O., Danchin, E., Mirouze, M., Loot, C. & Blanchet, S. 2016. Adaptation to global change: a transposable element-epigenetics perspective. *Trends Ecol. Evol.* **31**: 514-526.

Richards, C.L., Bossdorf, O. & Pigliucci, M. 2010. What role does heritable epigenetic variation play in phenotypic evolution? *BioScience* **60**: 232-237.

Richards, C.L., Schrey, A.W. & Pigliucci, M. 2012. Invasion of diverse habitats by few Japanese knotweed genotypes is correlated with epigenetic differentiation. *Ecol. Lett.* **15**: 1016-1025.

Richards, E.J. 2006. Inherited epigenetic variation-revisiting soft inheritance. *Nat. Rev. Genet.* **7**: 395-401.

Richards, E.J. 2008. Population epigenetics. *Curr. Opin. Genet. Dev.* **18**: 221-226.

Roberts, R.J., Vincze, T., Posfai, J. & Macelis, D. 2007. REBASE--enzymes and genes for DNA restriction and modification. *Nucleic Acids Res*. **35**: D269-270.

Roberts, S.B & Gavery, M.R. 2012. Is there a relationship between DNA methylation and phenotypic plasticity in invertebrates? *Front. Physiol*. **2**: 1-5.

Salathé, M., Van Cleve, J. & Feldman, M.W. 2009. Evolution of stochastic switching rates in asymmetric fitness landscapes. *Genetics* **182**: 1159-1164.

Salmon, A., Coltault, J., Jenczewski, E., Chable, V. & Manzanares-Dauleux, M.J. 2008. *Brassica oleracea* displays a high level of DNA methylation polymorphism. *Plant Sci.* **174**: 61-70.

Schield, D.R., Walsh, M.R., Card, D.C., Andrew, A.L., Adams, R.H. & Castoe, T.A. 2015. EpiRADseq: scalable analysis of genomewide patterns of methylation using next-generation sequencing. *Methods Ecol. Evol.* **7:** 60-69.

Schmitz, R.J., Schultz, M.D., Lewsey, M.G., O'Malley, R.C., Urich, M.A. & Libiger, O. *et al.* 2011. Transgenerational epigenetic instability is a source of novel methylation variants. *Science* **334**: 269-373.

Schrey, A.W., Coon, C.A.C., Grispo, M.T., Awad, M., Imboma, T. & McCoy E.D. *et al.* 2012. Epigenetic variation may compensate for decreased genetic variation with introductions: A case study using house sparrows *Passer domesticus* on two continents. *Genet. Res. Int.* **2012**: Article ID 979751.

Schrey, A.W., Alvarez, M., Foust, C.M., Kilvitis, H.J., Lee, J.D. & Liebl, A.L. *et al.* 2013. Ecological epigenetics: beyond MS-AFLP. *Integr. Comp. Biol.* **53**: 340-350.

Schübeler, D. 2015 Function and information content of DNA methylation. *Nature* **517**: 321-326.

Scott, G.R. & Johnston, I.A. 2016. Temperature during embryonic development has persistent effects on thermal acclimation capability in zebrafish. *Proc. Natl. Acad. Sci. USA* **109**: 14247-14252.

Seong, K., Li, D., Shimizu, H., Nakamura, R. & Ishii, S. 2011. Inheritance of stress-induced, ATF-2 dependent epigenetic change. *Cell* **145**: 1049-1061.

Shea, N., Pen, I. & Uller, T. 2011. Three epigenetic information channels and their different roles in evolution. *J. Evol. Biol.* **24**: 1178-1187.

Simpson, V.J., Johnson, T.E. & Hammen, R.F. 1986. *Caenorhabditis elegans* DNA does not contain 5-methylcytosine at any time during development or aging. *Nucleic Acids Res.* **14**: 6711-6719.

Skinner, M.K., Gurerrero-Bosagna, C., Haque, M.M., Nilsson, E.E., Koop, J.A.H. & Knutie, S.A. *et al.* 2014. Epigenetics and the evolution of Darwin's finches. *Genome Biol. Evol.* **6**: 1972-1989.

Slatkin, M. 2009. Epigenetic inheritance and the missing heritability problem. *Genetics* **182**: 845-850.

Smith, E.M., Hoi, J.T., Eissenberg, J.C., Shoemaker, J.D., Neckameyer, W.S. & Ilvarsonn, A.M. *et al.* 2007. Feeding *Drosophila* a biotin-deficient diet for multiple generations increases stress resistance and lifespan and alters gene expression and histone biotinylation patterns. *J. Nutr.* **137**: 2006-2012.

Smith, G., Smith, C., Kenny, J.G., Chaudhuri, R.R. & Ritchie, M.G. 2015. Genome-wide DNA methylation patterns in wild samples of two morphotypes of threespine stickleback. *Mol. Biol. Evol.* **32**: 888-895.

Smith, T.A., Martin, M.D., Nguyen, M. & Mendelson, T. 2016. Epigenetic divergence as a potential first step in darter speciation. *Mol. Ecol.* **25**: 1883-1894.

Soll, D.R., Morrow, B. & Srikantha, T. 1993. High-frequency phenotypic switching in *Candida albicans*. *Trends Genet.* **9**: 61-65.

Suzuki, M.M. & Bird, A. 2008. DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* **9**: 465-476.

Tal, O., Kisdi, E. & Jablonka, E. 2010. Epigenetic contribution to covariance between relatives. *Genetics* **184**: 1037-1050.

Taudt, A., Colomé-Tatché, M. & Johannes, F. 2016. Genetic sources of population epigenomic variation. *Nat. Rev. Genet.* **7**: 319-332.

Thurman, T.J. & Barrett, R.D.H. 2016. The genetic consequences of selection in natural populations. *Mol. Ecol.* **25**: 1429-1448.

Trucchi, E., Mazzarella, A.B., Gilfillan, G.D., Lorenzo, M.T., Schönswetter, P. & Paun O. 2016. BsRADseq: screening DNA methylation in natural populations of non-model species. *Mol. Ecol.* **25**: 1697-1713.

Tweedie, S., Charlton, J., Clark, V. & Bird, A. 1997. Methylation of genomes and genes at the invertebrate-vertebrate boundary. *Mol. Cell Biol.* **17**: 1469-1475.

Uller, T., English, S. & Pen, I. 2015. When is incomplete epigenetic resetting in gem cells favoured by natural selection? *Proc. R. Soc. B* **282**: 20150682

Valenzuela, N. & Lance, V. 2004. *Temperature-dependent sex determination in vertebrates.* Washington: Smithsonian Books.

Varriale, A. & Bernardi, G. 2006. DNA methylation and body temperature in fishes. *Gene* **385**: 111-121.

Vaugh, M.W., Tanurdzić, M., Lippman, Z., Jiang, H., Carrasquillo, R. & Rabinowicz, P.D. *et al.* 2007. Epigenetic natural variation in *Arabidopsis thaliana*. *PLoS Biol.* **5**: e174.

Veening, J., Stewart, E.J., Berngruber, T.W., Taddei, F., Kuipers, O.P. & Hamoen, L.W. 2008. Bet-hedging and epigenetic inheritance in bacterial cell development. *Proc. Natl. Acad. Sci. USA* **105:** 4393-4398.

Verhoeven, K.J.F., Jansen, J.J., van Dijk, P.J. & Biere A. 2010. Stress-induced DNA methylation changes and their heritability in asexual dandelions. *New Phytol.* **185**: 1108-1118.

Verhoeven, K.J.F. & Preite, V. 2014. Epigenetic variation in asexual reproducing organisms. *Evolution* **68**: 644-655.

Verhoeven, K.J.F., vonHoldt, B.M. & Sork, V.L. 2016. Epigenetics in ecology and evolution: what we know and what we need to know. *Mol. Ecol.* **25**: 1631-1638.

Via, S. & Lande, R. 1985. Genotype-environment interaction and the evolution of phenotypic plasticity. *Evolution* **29**: 505-522.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T. & Hornes, M. *et al.* 1995. AFLP-a new technique for DNA-fingerprinting. *Nucleic Acids Res.* **23**: 4407-4414.

Waddington, C.H. 1942. The epigenotype. *Endeavour* **1**: 18-20.

Wang, Y., Jorda, M., Jones, P., Maleszka, R., Ling, X. & Robertson, H.M. *et al.* 2006. Functional CpG methylation system in a social insect. *Science* **314**: 645-647.

Waterland, R.A. & Jirtle, R.L. 2003. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol. Cell. Biol.* **23**: 5293-5300.

Weaver, I.C., Cervoni, N., Champagne, F.A., D'Alessio, A.C., Sharma, S. & Secki, J.R. *et al.* 2004. Epigenetic programming by maternal behaviour. *Nature Neurosci.* **7:** 847-854.

Wenzel, M.A. & Piertney, S.B. 2014. Fine-scale population epigenetics structure in relation to gastrointestinal parasite load in red grouse (*Lagopus lagopus scotica*). *Mol. Ecol.* **23**: 4256-4273.

West-Eberhard, M.J. 2003. *Development Plasticity and Evolution*. Oxford University Press, New York.

Weyrich, A., Lenz, D., Jeschek, M., Chung, T.H., Rübensam, K. & Göritz, F. *et al.* 2016. Paternal intergenerational epigenetic response to heat exposure in male Wild guinea pigs. *Mol. Ecol.* **25**: 1729-1740.

Willing, E-M., Rawat, V., Mandakova, T., Maumus, F., James, G.V. & Nordström, K.J.V. *et al.* 2015. Genome expansion of *Arabis alpina* linked with retrotransposition and reduced symmetric DNA methylation. *Nat. Plants* **1**: 1-7.

Youngson, N.A. & Whitelaw, E. 2008. Transgenerational epigenetic effects. *Annu. Rev. Genomics Hum. Genet.* **9**: 233-257.

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

Zhang, Y.Y., Fischer, M., Colot, V. & Bossdorf, O. 2013. Epigenetic variation creates potential for evolution of plant phenotypic plasticity. *New Phytol.* **1**: 314-322.

Zhou, V.W., Goren, A. & Bernstein, B.E. 2011. Charting histone modifications and the functional organization of mammalian genomes. *Nat. Rev. Genet.* **12**: 7-18.

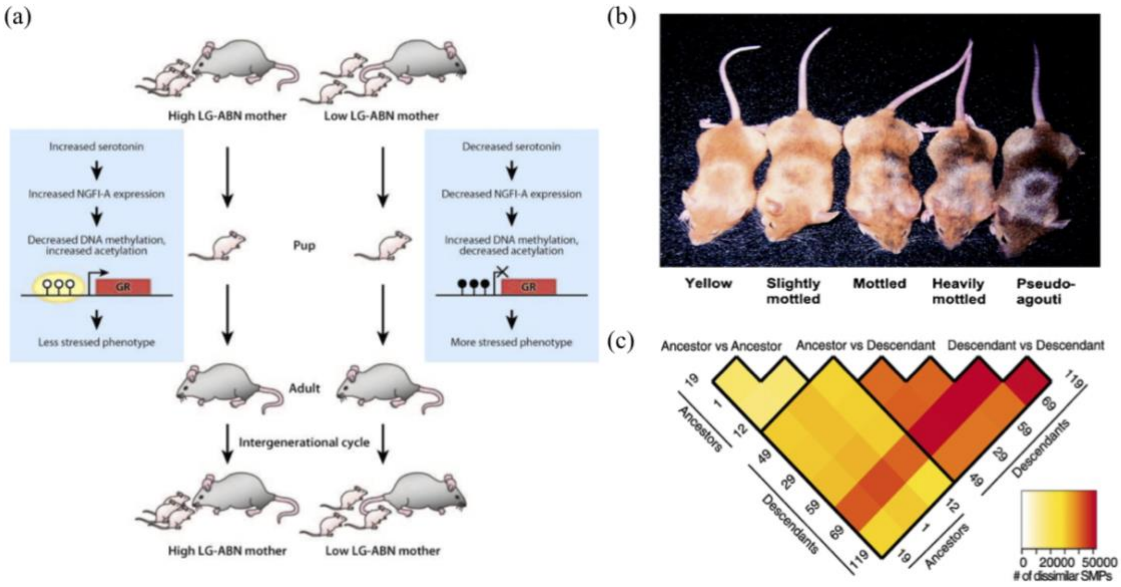


Fig. 1.1 Examples of roles of epigenetic variation in evolution. (a) Transgenerational inheritance of mothering style and stress in rat. Mothering style (licking/grooming (LG) and arched-back nursing (ABN)) that results in different DNA methylation and histone acetylation status at the promoter of glucocorticoid receptor (GR) gene provokes the occurrence of the same epigenetic markers in the offspring. (b) Maternal dietary methyl supplementation and coat colour phenotype of A^{vy}/a offspring. The methylation status of a transposable element at viable yellow agouti gene (A^{vy}) controls coat colours of isogenic A^{vy} /a mice. The A^{vy} alleles of yellow mice (left) are hypomethylated, allowing maximal ectopic agouti expression. A^{vy} hypermethylation silences ectopic agouti expression in pseudoagouti animals, recapitulating the agouti phenotype (right). (c) A heatmap indicating the number of CG single methylation polymorphisms (CG-SMPs) that differ between ancestral and descendant *Arabidopsis* populations. Although the total number of CG-SMPs was similar between all lines, the conservation of these polymorphisms among and between ancestral and descendant populations was different. Reproduced with permission from (a) Youngson & Whitelaw (b) Waterland & Jirtle and (c) Schmitz *et al*.

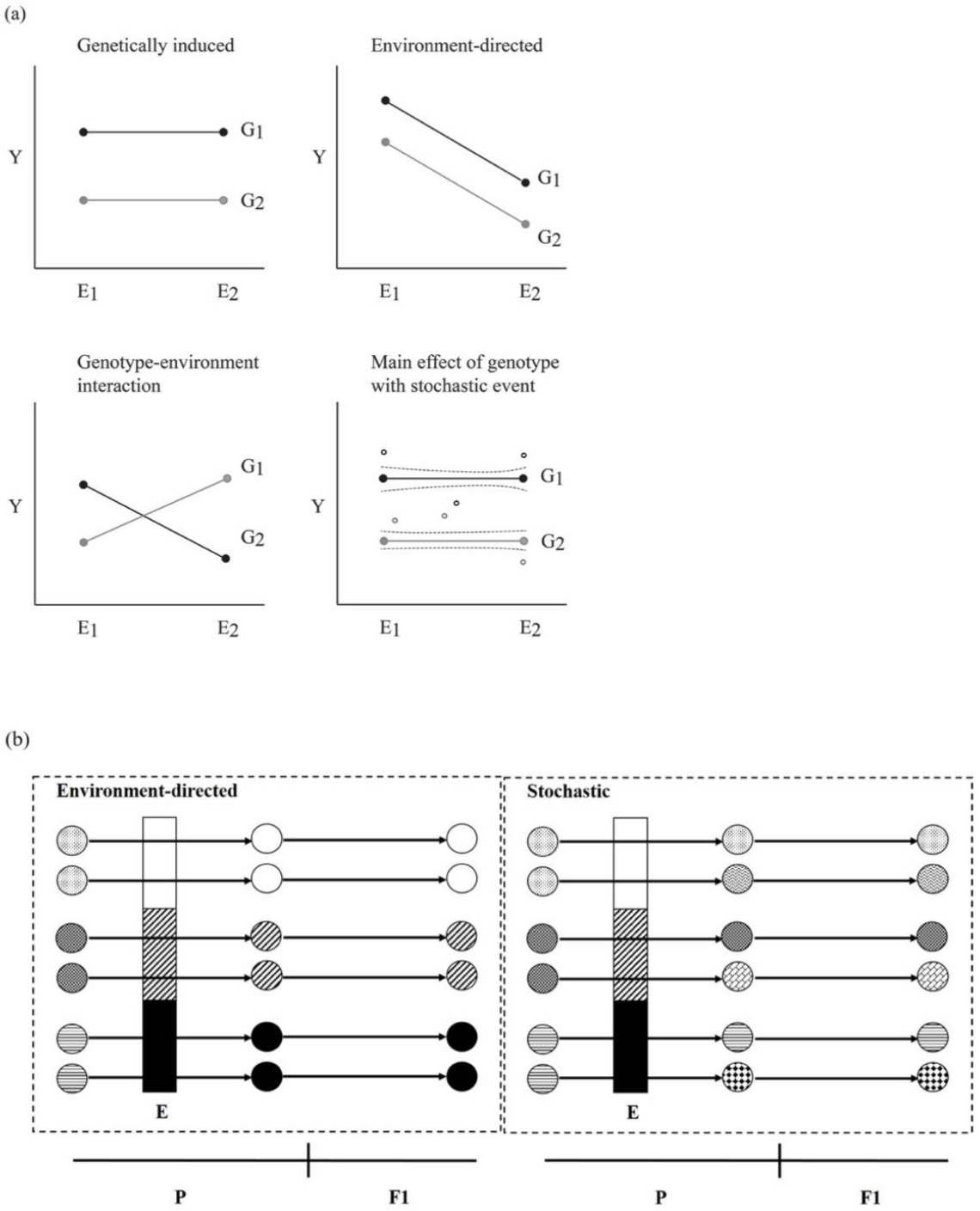


Fig. 1.2 A schematic of the source and stability of epigenetic variation during adaptation at the population level. (a) Interaction plots of different main sources of epigenetic variation. Here we assume a scenario with two genotypes (G1 and G2) and two environments (E1 and E2). The Y-axis plots the chromatin state. Lines connect means of each genotype in each environment. We show 95% confidence interval around means. (b) Relationships between environment-directed epigenetic variation, stochastic epigenetic variation, and environmental change. Here we assume a starting population (P) with epigenetic variation among

individuals (dot, hatched, and lined circles). P individuals respond to environmental change (E), and produce epigenotypes that are stably transmitted to the F1 generation. In the scenario of environment-directed epigenetic variation, epigenotypes are produced based on the cues that P individuals experience from the environment (indicated by matching of epigenotype pattern and temperature pattern), and the amount of epigenetic variation remains constant. In the scenario of stochastic epigenetic variation, new epigenotypes are produced at random and without regard to environmental conditions (wave, brick, and checker board), and thus the epigenetic variation is increased.

Table 1.1 Overview of studies demonstrating epigenetically encoded plasticity in animals

Animal species	Phenotype(s)	Epigenetic modification	Generations assayed (effect detected)	Richards' framework	Ref.
European sea bass (*Dicentrarchus labrax*)	Sex ratio	DNA methylation	1 (1)	Putatively obligatory (*cis*)	Navarro-Martin *et al.* (2011)
Red-ear slider turtle *(Trachemys scripta)*	Sex ratio	DNA methylation	1 (1)	Putatively obligatory (*cis*)	Matsumoto *et al.* (2013)
American alligator (*Alligator mississippiensis*)	Sex ratio	DNA methylation	1 (1)	Putatively obligatory (*cis*)	Parrott *et al.* (2014)
Senegalese sole (*Solea senegalensis*)	Muscle fibre diameter	DNA methylation	1 (1)	Putatively obligatory (*trans*)	Campos *et al.* (2013)
Zebrafish (*Danio rerio*)	Muscle phenotype (hyperplastic vs. hypertrophic)	microRNA	1 (1)	Putatively obligatory (*trans*)	Johnston *et al.* (2009)
Cobb chick (*Gallus gallus domesticus*)	Thermotolerance acquisition	Histone modification	1 (1)	Putatively obligatory (*trans*)	Kisliouk & Meiri (2009)
Cobb chick (*Gallus gallus domesticus*)	Thermotolerance acquisition	microRNA	1 (1)	Putatively obligatory (*trans*)	Kisliouk *et al.* (2011)
Antarctic polychaete (*Spiophanes tcherniai*)	Metabolic rates	DNA methylation	1 (1)	Putatively obligatory (*trans*)	Marsh & Pasqualone (2014)
Carp (*Cyprinus carpio*)	Nucleolar organization	Chromatin structure	1 (1)	Unknown	Alvarez *et al.* (2006)
Fruit fly (*Drosophila melanogaster*)	Life span, heat tolerance	Histone modification	1 (1)	Putatively obligatory (*trans*)	Camporeale *et al.* (2006)
	Life span, heat tolerance, fertility, metabolism		1 (1)	Putatively obligatory (*trans*)	Smith *et al.* (2007)
	Eye colour	Chromatin structure	5 (2)	Putatively obligatory (*cis* and *trans*)	Seong *et al.* (2011)

Guinea pigs (*Cavia aperea*)	None	DNA methylation	2 (2)	Both obligatory and facilitated/pure	Weyrich *et al.* (2016)
Reef corals (A*cropora hyacinthus, A. millepora, A. palmate, Pocillopora damicornis, Porites astreoides, Stylophora pistillata*)	None	DNA methylation	2 (possible 2)	Obligatory (*cis*)	Dimond & Roberts (2016)

Table 1.2 Models of heritable epigenetic variation and evolution

Approach	Relationship with genetic variation	Changing environment?	Effects of epigenetic variation	Example Theor. refs	Example Empir. refs
Effects of switching between epigenetic variants on phenotypes	Not mentioned	Yes	Long-term inheritance of epigenetic variants result in reduced fitness variance and greater population growth in random and temporally patchy environment.	Jablonka *et al.* (1995)	None
	Not mentioned	Yes	In an environment that both induces and selects variants, epigenetic variants that have a switching rate corresponding to environmental periodicity will be advantageous.	Lachmann & Jablonka (1996)	Soll *et al.* (1993)
	Controlled by genetic variation	Yes	Switching rate of epigenetic variation will reflect the rate of environmental change when selection is strong.	Salathé *et al.* (2009)	None
	Controlled by genetic variation	Yes	Increased phenotypic variability and fitness in changing environments, without changes in the mean phenotype.	Feinberg & Irizarry, (2010)	Feinberg & Irizarry (2010)
	Independent from genetic variation	Yes	In the short term, epigenetic changes are equivalent to mutations and are likely to be in LD with SNPs. In the long term, epigenetic changes have limited contribution to heritability and recurrence risk in disease.	Slatkin (2009)	None
	Controlled by genetic variation	Yes	An equilibrium exists for the frequencies of the alleles at the locus controlling the epigenetic states. The equilibrium depends on the mutation rate, the fitness landscape, and the period of the environmental fluctuation.	Carja & Feldman (2012)	None
	Not mentioned	Yes	Maintained phenotypic variation in changing environments, even without genetic variation.	Geoghegan & Spencer (2012, 2013)	Herrera *et al.* (2013)
	Controlled by genetic variation	Yes	Serve as a mechanism for plasticity, phenotypic switching, or stable inheritance of phenotypic states, depending on the	Furrow & Feldman (2014)	None

			frequency of environmental changes.		
	Not mentioned	Yes	Transmission of epigenetic states prevents mismatched phenotypes when the environment changes infrequently relative to generation time and when maternal and environmental cues are unreliable.	Uller *et al.* (2015)	None
	Controlled by genetic variation	Yes	Selection can favour epigenetic variation that generates a positive correlation between parental and offspring phenotype under relatively stable environments, and generates a negative correlation between parental and offspring phenotype under unstable environments.	Kuijper & Johnstone (2016)	Lewis (2010); Schmitz *et al.* (2011)
Effects of epigenetic mutation on maintaining genetic or phenotypic variation	Controlled by genetic variation	No, but with changes in microenvironments	Promote population persistence, particularly when the mean phenotype is far from the optimum. Genetic assimilation occurs much closer to the fitness peak.	Pál (1998); Pál & Miklos (1999)	Ho *et al.* (1983)
	Not mentioned	Yes	Determine phenotypic variation, and decouple phenotypic change from evolutionary dynamics of genotype.	Day & Bonduriansky, 2011	Skinner *et al.*, 2014
	Independent of genetic variation	Yes	Decouple fitness from genetic variation, allow populations to have higher level of genetic variation, and allow populations to respond to environmental change even without genetic variation.	Klironomos *et al.* (2013)	None
	Independent of genetic variation	No	Depending on stability and fitness effects relative to genetic variation, epigenetic variation can accelerate the initial stage of adaptation, but cause lower final fitness values, or vice versa.	Kronholm & Collins (2015)	None
	Not mentioned	Yes	The combination of epigenetic inheritance and developmental noise can also explain some bet-hedging strategies that were previously explained solely by genetic mechanisms	Day (2016)	None

CONNECTING STATEMENT

In Chapter 1, I reviewed recent empirical and theoretical studies analysing the role of epigenetics in phenotypic plasticity and evolution. Although it is evident that epigenetics can contribute to phenotypic change, local adaptation, and evolutionary processes, results have typically come from studies using lab-raised animals and have not explicitly considered ecological context. To add to the growing body of literature on epigenetics in natural populations of animals, I evaluate epigenetic responses to environmental conditions in three distinct empirical systems. In Chapter 2, I investigate changes in genome-wide DNA methylation patterns of Trinidadian guppies (*Poecilia reticulata*) during the course of infection by the monogenean ectoparasite, *Gyrodactylus turnbulli*.

CHAPTER 2: Genome-wide DNA methylation signatures to infection status in Trinidadian guppies (*Poecilia reticulata*)

Authors: JUNTAO HU,* FELIPE PÉREZ-JVOSTOV,* LÉA BLONDEL*, and ROWAN D. H. BARRETT*

Initials: J.H., F.P.J., L.B. and R.D.H.B.

Keywords: DNA methylation, epigenetics, host-parasite interactions, phenotypic plasticity, reduced representation bisulphite sequencing, *Gyrodactylus turnbulli*, *Poecilia reticulata*

*Redpath Museum and Department of Biology, McGill University, 859 Sherbrooke Street West, Montreal, Quebec, Canada H3A 0C4

Correspondence: Juntao Hu, Fax: +1 514 398 3185; E-mail: juntao.hu@mail.mcgill.ca

Abstract

Epigenetic modification, especially DNA methylation, can play an important role in mediating gene regulatory response to environmental stressors, and may be a key process affecting phenotypic plasticity and adaptation. Parasites are potent stressors with profound physiological and ecological effects on their hosts, yet it remains unclear how parasites influence host methylation patterns. Here, we used a well-studied host-parasite system, the guppy *Poecilia reticulata* and its ectoparasitic monogenean *Gyrodactylus turnbulli* to gain mechanistic insight into the dynamics of DNA methylation in host-parasite interactions. To explore this, we quantitatively measured genome-wide DNA methylation in guppy skin tissue using reduced representation bisulphite sequencing, and characterised differential methylation patterns in guppies during distinct phases of infection. We identified 365, 313, and 741 differentially methylated regions (DMRs) between *infected* and *control* fish in *early infection*, *peak infection,* and *recovery* phases, respectively. The magnitude of the methylation difference was moderate in DMRs, with an average of 29% (*early infection*), 27% (*peak infection*), and 30% (*recovery*) differential methylation per DMR. Approximately 50% of DMRs overlapped with CpG islands, and over half of the DMRs overlapped with gene bodies, several of which encode proteins relevant to immune response. These findings provide the first evidence of an epigenetic signature of infection by ectoparasites, and demonstrate the changing relationship between epigenetic variation and immune response in distinct phases of infection.

Introduction

Parasitism has long been recognised as a major driver of ecological and evolutionary processes in a wide range of host taxa (Hamilton 1980; Sheldon & Verhulst 1996; Paterson & Piertney 2011). This relevance is classically attributed to increased mortality of infected individuals, and to parasite-induced changes in host phenotype (Hatcher *et al.* 2006). Examples of both mechanisms are plentiful, both in the laboratory (e.g., Lazzaro *et al.* 2008; Hari Dass & Vyas 2014) and in the wild (e.g., van Oosterhout *et al.* 2003; Gotanda *et al.* 2013), and their ecological and evolutionary implications have been well documented (Penczykowski *et al.* 2016). Recently, the role of epigenetic modulation in host-parasite interactions has received increased attention as a potential source of rapid and reversible phenotypic variation that can be shaped by both parasites and the host (Gómez-Díaz *et al.* 2012; Silmon de Monerri & Kim 2014; Cheeseman & Weitzman 2015; Robert McMaster *et al.* 2016). Indeed, parasites can modulate gene expression profiles in their hosts through epigenetic modifications (Paschos & Allday 2010; Sessions *et al.* 2013), and these modifications can also be associated with an adaptive immune response of the host (Youngblood *et al.* 2010; Boyko & Kovalchuk 2011; Conrath 2011; Holeski *et al.* 2012). However, to date, the most compelling evidence for epigenetic responses to parasite infection has come from studies of endoparasites, such as bacterial pathogens infecting plants (e.g., Dowen *et al.* 2012) and intracellular protozoans infecting vertebrates (e.g., Hari Dass & Vyas 2014). To our knowledge, no study has investigated epigenetic responses of hosts to ectoparasites, or if these responses change during the course of an infection. Ectoparasites are distinct from endoparasites in that they cannot manipulate host cell machinery, and thus cannot directly modify intracellular signalling pathways and host transcription regulation (Cheeseman & Weitzman 2015). Thus, the effects that ectoparasites have on host epigenome are unknown, but they are likely to be different from those of intracellular parasites. Here, we explore epigenetic modifications in a host-parasite system that is ideal for testing dynamics across distinct phases of infection, and that has been extensively studied both in nature and in the laboratory: Trinidadian guppies and their monogenean ectoparasites, *Gyrodactylus*.

Trinidadian guppies have been frequently used in evolutionary studies due to their dramatic and rapid adaptation to the local environment (for reviews: Endler 1995; Houde 1997; Magurran 2005). Although initial work mainly focused on interactions between guppies and their predators (Reznick & Endler 1982; Reznick *et al.* 1997), an extensive body of work has shown that gyrodactylid ectoparasites can influence many aspects of guppy ecology, including mate choice (Kennedy *et al.* 1987; López 1999), foraging behaviour

(Kolluru *et al.* 2006), life-history traits (Pérez-Jvostov *et al.* 2012, 2017), bacterial and fungal infections (Cusack & Cone 1986; Kotob *et al.* 2016), and survival (van Oosterhout *et al.* 2007; Pérez-Jvostov *et al.* 2017). *Gyrodactylus* spp. are skin grazing parasites with a hyperviviparous life cycle, where the first-born offspring develops asexually from the adult female, and contains a developing embryo (Bakke *et al.* 2007). A single adult *Gyrodactylus* worm can produce rapid exponential population growth on individual fishes, and such infection has been shown to have significant fitness consequences (Bakke *et al.* 2007), for example, infected fish have higher mortality (van Oosterhout *et al.* 2007), and lower mating rates (Kennedy *et al.* 1987). Thus, studying the epigenetic mechanisms underlying responses to *Gyrodactylus* in guppies will aid in understanding a wide range of ecological and evolutionary processes in this well-studied host-parasite system.

As a first defence against skin grazing parasites like *Gyrodactylus*, fish largely rely on their innate immune system in the form of localised inflammation, which typically appears as hyperplasia, and elevated mucus secretion on the skin (Kumar *et al.* 2017). Increased expression of cyto- and chemokines in fish skin has been observed on infected fish mounting an immune response against a *Gyrodactylus* spp. infection (Lindenstrøm *et al.* 2003; Lindenstrøm *et al.* 2004; Matejusová *et al.* 2006; Kania *et al.* 2007). Similarly, skin mucus contains high concentrations of lectins and immunoglobulins that also play important roles in both detecting and attacking ectoparasites (Salinas *et al.* 2011; Ángeles Esteban 2012). Nonetheless, *Gyrodactylus* can often overcome these defenses and reach such infection levels (hundreds of worms on one fish) that their grazing on the skin can diminish mucus production, and quickly decrease host health (Wells & Cone 1990; Buchmann & Bresciani 1997). Taken together, it is evident that *Gyrodactylus* infections can change the physiology of infected fish, and can also result in changes in gene expression. However, the role of epigenetic mechanisms in regulating these changes is unknown.

We address this gap by performing an extensive epigenomic survey of guppies during three distinct phases of infection with a guppy-specific *Gyrodactylus*. Recent studies have shown increased genome-wide methylation as a general response to infection by endoparasites (Paschos & Allday 2010; Hari Dass & Vyas 2014; Marr *et al.* 2014). One explanation for this pattern is that parasites 'hijack' the epigenome of the host by inducing hypermethylation in promoters of immune genes, which can result in gene repression and thus allow parasites to evade host defence mechanisms (Silmon de Monerri & Kim 2014). In contrast, we predict that ectoparasites such as *Gyrodactylus* should not be able to manipulate the epigenome of guppies in this way, and thus we have the opposite expectation: an active

immune response by guppies should be reflected by increased expression of immune genes and thus hypomethylation of their promoters. To test this hypothesis, we specifically surveyed methylation changes over distinct phases of infection. Recent work has shown that host gene expression can vary across the course of infection, thereby indicating functional changes in immune response at different points of an infection cycle (Choi *et al.* 2014; Westermann *et al.* 2017). If methylation responses to infection reflect an active immune response of guppies to their parasites, we expect to see the greatest methylation changes occurring in fish that are able to successfully recover from infection. We investigate three specific questions: (1) Are there general methylation patterns associated with *Gyrodactylus* infection in the guppy genome? (2) How are differentially methylated regions distributed among different regions of the genome (i.e., promoters, exons, introns, and intergenic regions)? (3) How do methylation patterns change during the course of infection? Answering these questions will help us to better understand the mechanisms underlying host responses to parasite infection at the molecular level.

Materials and Methods

Guppies

We used females from an admixed, lab-reared, population of guppies sampled from different locations in throughout the Northern Mountain range in Trinidad that have been in the laboratory at McGill University for at least ten generations. We selected only females because they have a broader range of peak *Gyrodactylus* loads (Cable & van Oosterhout 2007), and because methylation patterns are known to be sex-specific (McCarthy *et al.* 2014). Prior the initiation of the experiment, we selected over 60 guppies from this lab population, and scanned them under the microscope to identify if *Gyrodactylus* was prevalent. From these fish, only one male was infected, and it had only one worm. Although this prevalence is very low, we do not know the history of *Gyrodactylus* infections in this lab-reared population, and it is thus possible that some fish have previously been infected. Therefore, we randomly selected healthy females, treated them for *Gyrodactylus* using Clout (Fritz Industries Inc., Mesquite, TX, USA) – in case we missed any worms – and kept them isolated for a three-week quarantine period. We expect that after this quarantine, any females with prior exposure to *Gyrodactylus* would have lost any acquired immune response to *Gyrodactylus*, and respond to infection as naïve fish (Scott 1985; Richards & Chubb 1996). Fish were monitored daily to ensure that they were in good health, and were fed daily with brine shrimp. At the

end of the three-week quarantine period all fish were confirmed to be free of *Gyrodactylus* infections.

Experimental infections

Prior to the infection trials all quarantined females were transferred to individual 1.8L tanks in a flow through system (Aquaneering Inc., San Diego, CA, USA) that standardises water quality and temperature (26°C), and prevents potential movement of parasites between tanks with ultraviolet sterilizers. Females were divided into two experimental groups: *control* and *infected* (see Experimental design below; Fig. 2.1). To initiate the infections in females in the *infected* group, each female was anaesthetised using MS-222 (buffered to a neutral pH with $NaHCO_3$), and manually infected by transfering two to three *Gyrodactylus* worms from one donor infected guppy (day 0). Females in the *control* group underwent the same procedure, except they were sham infected. For all trials we used an isogenic strain of a guppy-specific *Gyrodactylus, G. turnbulli*, that was isolated from one worm from a pet store guppy in Montreal in 2008 (Dargent *et al.* 2013; Tadiri *et al.* 2013), and has been kept at high densities in the lab using pet store guppies as hosts (Tadiri *et al.* 2016). Given that neither this isogenic strain of *Gyrodactylus* nor the guppies used in this experiment have had prior exposure to each other, our female guppies are naïve to this specific strain of *G. turnbulli* (hereafter refered to simply as *Gyrodactylus*).

After the experimental infection, all *control* and *infected* females were scanned for *Gyrodactylus* under a dissecting microscope every two days to track the development of the infection, or to confirm the absence of it in the control. This is a standard procedure to count *Gyrodactylus* worms, and does not affect guppy health (Scott 1982, 1985; Pérez-Jvostov *et al.* 2012; Dargent *et al.* 2013).

Experimental design

We designed our experiment with two main factors: experimental group (*infected* and *control*), and phase (*early infection*, *peak infection*, and *recovery*; Fig. 2.1). Given the high variability in parasite numbers at any specific time post infection, we controlled for infection intensity rather than days post infection. Following Gheorghiu *et al.* (2012), we characterised an *early infection* phase in which guppies had parasite loads < 20 worms, and a *peak infection* phase in which guppies had ~100 parasites and/or showed erratic swimming or decreased health. We also characterised a *recovery* phase in which, after initial *Gyrodactylus* population

growth, the number of worms on the fish started to decrease for two consecutive days or to half the number of the previous scan, suggesting guppy immune response was decreasing *Gyrodactylus*' population growth. We recorded the day post infection when the number of parasites in *infected* fish reached each of the three infection phases, and found similar infection dynamics of *infected* fish within each infection phase, with the exception of one fish from *peak infection* phase, suggesting that *infected* fish had similar temporal responses to infection (Fig. 2.1). We collected the complementary *control* fish on the same day when we collected *infected* fish to control for their age. In total, we collected three *infected* fish from each of the *early* and *peak infection* phases, as well as three complementary *control* fish. For the *recovery* phase we were able to only sample three *infected* and two *control* fish.

DNA extraction

We extracted DNA from epidermis of skin using phenol:chloroform:isoamyl alcohol (25:24:1), and assessed the quality and quantity using Tecan Infinite® 200 NanoQuant and Quant-iT PicoGreen® dsDNA assay kit (ThermoFisher Scientific). We used skin tissue because it is in intimate contact with the parasite, and involved in immune response to ectoparasites in fish (Ángeles Esteban 2012). All procedures were approved by McGill University (Animal Use Protocol 2000-4570).

Reduced representation bisulphite sequencing

To measure genome-wide DNA methylation levels, we used a high-throughput sequencing approach known as reduced representation bisulphite sequencing (RRBS) (Meissner *et al.* 2008; Gu *et al.* 2011), following the protocol described in Boyle *et al.* (2012) with minor modifications. For each individual, we created a library from 200 ng of genomic DNA, and ligated the fragments in each library with unique Illumina TruSeq adapters. We targeted fragments of 160-340bp (including ~120bp of adapter sequence) using NaCl-PEG diluted SpeedBeads (Rohland & Reich 2012). We randomly multiplexed eight and nine libraries into two pools, and treated the pools with sodium bisulphite (EpiTect, Qiagen), following the protocol for formalin-fixed paraffin-embedded samples. After two rounds of bisulphite treatment to ensure complete conversion, each pool was amplified with Illumina primers, and loaded in two lanes (100-bp single-end reads) of Hiseq 2500 at the McGill University and Genome Quebec Innovation Centre. Each sample was sequenced to a mean depth (±SD) of 19.86 ± 4.210 million reads (Table S2.1, Supporting information). We quantified methylation

at non-CpG motifs, and found less than 1% non-CpG cytosines were methylated, suggesting a highly efficient bisulphite conversion.

Read filtering and mapping

To remove adapter contamination, low-quality bases, and bases artificially introduced during library construction, we trimmed reads using Trim Galore! v0.4.4 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), with the 'rrbs'option. We then used the program Bowtie2 v2.2.9 (Langmead & Salzberg 2012), implemented in Bismark v0.17.0 (Krueger & Andrews 2011) to align trimmed reads for each sample to the guppy genome (GenBank assembly accession GCA_00063615.2) with default settings, except for tolerating one non-bisulphite mismatch per read. We only included reads that mapped uniquely to the reference genome in downstream differentially methylated cytosines (DMCs) and differentially methylated regions (DMRs). The average mapping efficiency (± SD) was 58.69 ± 2.03%. Only CpG context cytosine methylation was analysed because CpG methylation is the most common functional methylation in vertebrates (Suzuki & Bird 2008).

DMC and DMR calling

We analysed DMCs and DMRs at two steps: first, we pooled all *infected* or *control* fish as two groups, and identified DMCs and DMRs between groups to determine general patterns of methylation response to parasites in the guppy genome. Second, we performed similar DMC and DMR analyses within each infection phase to examine specific epigenetic responses in each phase of infection. When pooling all *infected* or *control* fish, a total of 878,645 CpG sites met the minimum coverage requirement, consisting of an average of 1.5% of all CpG sites in the genome after alignment (Table S1, Supporting information). For each phase, 978,671 (*early infection*), 1,018,862 (*peak infection*), and 1,027,149 (*recovery*) CpG sites met the minimum coverage requirement, consisting of ~2 % of all CpG sites after alignment (Table S1, Supporting information).

We identified individual DMCs using the R package methylKit v1.4.1 (Akalin *et al.* 2012). Read coverage was normalized between samples. A minimum of five reads in all samples were required at a CpG site for that site to be analysed (Walker *et al.* 2015; Wan *et al.* 2016). Sites that were in the 99.9th percentile of coverage were also removed from the analysis to account for potential PCR bias. Hierarchical cluster analysis was conducted using Ward's method based on the filtered CpG sites, and most individuals clustered primarily by

experimental group (*infected* vs. *control*; Fig. S2.1). We used default parameters (false discovery rate correction Q-value < 0.01), with a correction for overdispersion, and a minimum required methylation difference of 25% between *infected* and *control* fish to identify DMCs (Akalin *et al.* 2012; Baerwald *et al.* 2015). We then determined DMRs using the R package eDMR v0.6.4.1 (Li *et al.* 2013) with default parameters. To be considered significant, a DMR needed to contain at least three CpG sites within an algorithm-specified genomic distance, with at least one classified as a DMC (Q-value < 0.01), and an absolute mean methylation difference greater than 20% when comparing *infected* and *control* fish (see Li *et al.* 2013 for details). We analysed the shared DMRs between different infection phases by extracting and comparing the chromosomal names, and the start and end positions of each DMR. We visualised differential methylation patterns across individuals, and obtained clustering of samples and DMRs in heatmaps with the "complete" clustering method on Euclidian distances, using the R package pheatmap v1.0.8 (https://cran.r-project.org/web/packages/pheatmap/index.html). We clustered hyper- and hypomethylated DMRs between *infected* and *control* fish using the relative percent DNA methylation, which is the normalised percent DNA methylation scaled for each DMR's percent DNA methylation (median per cent methylation as 0) of *infected* and *control* fish in heatmaps. We also clustered individual fish based on overall methylation patterns across DMRs.

Genomic context, gene annotation, and gene ontology analysis

To build null distributions of genomic features (promoters/exons/introns/intergenic regions) of DMCs, we used the sets of CpG sites that passed the filtering steps described above. We first identified the positions of CpG sites within genomic features. We gave precedence to promoters > exons > introns > intergenic regions when features overlapped, and defined promoter regions as upstream 1000 bp and downstream 1000 bp from the transcription starting site (TSS) (Akalin *et al.* 2015). We then identified the positions of DMCs within genomic features, and compared the distributions of DMCs to null distributions using G-tests.

To perform the functional analysis of DMRs, we identified the nearest TSS and its associated transcript ID, and the position of methylated regions within genomic features for each DMR. We also identified the proximity of DMRs to CpG islands, which are CpG-rich regions that are usually unmethylated and serve as sites for transcription initiation (Jones 2012), using python scripts (https://github.com/lucasnell/TaJoCGI) that apply an algorithm based on the methods described in Takai & Jones (2002). DMRs were considered

overlapping or proximal to CpG islands when they were within or less than 4 kb away from an island, and considered locating within open sea when they were outside the 4 kb window (Baerwald *et al.* 2015). As the guppy genome has not been fully annotated yet, we used BLASTx against the NCBI non-redundant database to identify genes that DMRs were mapped to, followed by functional category assignment, GO term mapping, and node score distribution analysis implemented in Blast2GO v4.1 (Conesa *et al.* 2005; Götz *et al.* 2008). To specifically investigate the immune relevance of identified genes, we acquired a list of 1,843 GO terms including 'immune system process' (GO:0002376) and its child terms using the R package GO.db v3.4.0 (Carlson 2017), and compared the blast results of DMR-associated transcripts to the list. We then created a list of DMRs overlapping with genes or regions up to 5 kb upstream or downstream of these gene locations (Le Luyer *et al.* 2017). In addition, we also checked if any of the DMRs overlapped with the promoters of immune genes, using the annotatePeakInBatch function implemented in the R package ChIPpeakAnno v3.14.0 (Zhu *et al.* 2010; Zhu 2013).

Pathway analysis

To identify functional associations among the genes that DMRs were mapped to, we conducted pathway analysis using the Kyoto Encyclopaedia of Genes and Genomes (KEGG) enrichment analysis implemented in Blast2GO v4.1 (Conesa *et al.* 2005; Götz *et al.* 2008).

Results

General patterns of differential methylation in infected vs control guppies

After pooling across phases (nine *infected* fish vs. eight *control* fish), we identified 1,540 DMCs between *infected* and *control* fish after false discovery rate correction. We found significantly more DMCs within exons (*G*-test; $P = 0.028$) than expected by chance, and fewer within promoters (*G*-test; $P < 0.001$) when compared to the null distribution built on all CpG sites that passed the filtering steps; however, DMCs were not significantly enriched in introns (*G*-test; $P = 0.49$) or intergenic regions (*G*-test; $P = 0.36$; Fig. 2.2a). Based on the DMCs, we found 30 DMRs between *infected* and *control* fish after false discovery rate correction. Given our lower limit differential methylation cut-off (20%), methylation differences between *infected* and *control* group ranged from 20 to 45% per DMR (Table S2.2, Supporting information). Based on Euclidean distances calculated from the 30 DMRs (Fig. 2.2b), most individuals clustered primarily by experimental group (*infected* vs. *control*). Both

DMCs and DMRs displayed more hyper- than hypomethylation in *infected* fish relative to *control* fish (872 hypermethylated and 668 hypomethylated DMCs, 18 hypermethylated and 12 hypomethylated DMRs when comparing *infected* fish to *control* fish), suggesting that an increase in genomic DNA methylation levels is a general response to parasite infection in guppies.

We mapped 27 of the 30 DMRs to 15 chromosomes, with 25 of the DMRs localised within or proximal to known genes. Chromosome LG10 (NC_024340.1) contained the most DMRs (four). Four other chromosomes, LG6, LG13, LG14, and LG16 (NC_024336.1, NC_024343.1, NC_024344.1 and NC_024346.1) contained two to three DMRs. Ten chromosomes each contained a single DMR, and three DMRs could not be mapped to a chromosome due to the incomplete nature of the guppy genome. Over half of the DMRs were found overlapping with gene bodies, and 30% of the DMRs (10/33) were found within or proximal to CpG islands. The DMRs mapped to genes annotated with a variety of gene ontology (GO) categories in Biological Process, and the gene ontology categories with the highest node scores found were ATP binding (GO:0005524) and zinc ion binding (GO:0008270) in Molecular Function, and integral component of membrane (GO:0016021) in Cellular Component (Fig. S2.2, Supporting information). One DMR mapped to a gene (*ifngr1*) annotated with immune response (Table 2.1).

Patterns of differential methylation between the three distinct infection phases

After correcting for false discovery rates, we identified 11,355, 9,310, and 19,058 DMCs in *early infection*, *peak infection*, and *recovery* respectively. DMCs were distributed broadly across the genome, with no apparent clustering on specific chromosomes or chromosomal regions (Fig. 2.3a-c). More DMCs were hypermethylated than were hypomethylated in all phases (Fig. 2.3d), suggesting that the increase in genomic DNA methylation levels is a consistent response throughout all phases of infection. While we found significantly fewer DMCs within promoters than expected by chance in all phases (*G*-test; $P < 0.001$), DMCs were not significantly enriched in exons, introns, or intergenic regions in all phases (Fig. 2.4a-c). Based on DMCs, we identified 365 (*early infection*), 313 (*peak infection*), and 741 (*recovery*) DMRs, with most DMRs (~90%) located within or proximal to annotated genes in all phases (Fig. 2.4d-f). Over half of the DMRs (56% in *early infection* phase; 52% in *peak infection* phase; 57% in *recovery* phase) were found overlapping with gene bodies (Fig. 2.5a), and ~50% of the DMRs (47% in *early infection* phase; 46% in *peak infection* phase; 47% in

recovery phase) were found within or proximal to CpG islands (Fig. 2.5b). Methylation differences per DMR between *infected* and *control* groups ranged from 20 to 56% (*early infection*), 20 to 55 % (*peak infection*), and 20 to 70% (*recovery*). Among all identified DMRs, 26, 56, and 38 DMRs were shared between *early* and *peak infection* phases, *peak infection* and *recovery phases*, and *early infection* and *recovery phases*, respectively. Seven DMRs, which were annotated with known genes (*ADCY8*, *ANXA5*, *ARFGEF3*, *LRFN5, PPKG1, PPM1B,* and *PSMC3*), were shared by all three phases (Fig. 2.6a).

About 8% of all DMRs in our study were not located within or proximal to known genes. This may be an indication of *trans*-acting regulatory elements, e.g., enhancers (Taudt *et al.* 2016), but could also be due to the incomplete annotation of the guppy genome, precluding identification of genes, and their proximal regulatory regions. Gene ontology categories with the highest node scores were similar in all phases. These included single-organism cellular process (GO:0008150) and signal transduction (GO: 0007165) in Biological Process, hydrolase activity (GO: 0016787), nucleic acid binding (GO: 0003676) and transferase activity (GO: 0016740) in Molecular Function, and integral component of membrane (GO:0016021) in Cellular Component (Fig. S2.3-S2.5, Supporting information). We found seven, five, and seven DMRs annotated with immune response in *early infection*, *peak infection*, and *recovery* phases respectively (Table 2.1). None of the DMRs overlapped with the promoters of immune genes, and none of the immune-related GO terms were shared by all individual phases (Fig. 2.6b).

Pathway analysis

We identified several molecular pathways associated with *Gyrodactylus* infection for each infection phase via KEGG analysis (Table S2.3-S2.5, Supporting information). The top canonical pathway was purine metabolism for all three infection phases. Other top canonical pathways included aminobenzoate degradation, thiamine metabolism, Th1 and Th2 cell differentiation, and the T cell receptor signaling pathway.

Discussion

The modulation of host-parasite interactions through epigenetic mechanisms has received increased attention in recent years (Poulin & Thomas 2008; Gómez-Díaz *et al.* 2012; Wenzel & Piertney 2014). However, little is known about the role of epigenetics in host responses to ectoparasites, or if and how these responses change throughout the duration of an infection.

We used a quantitative, single-base resolution technique (RRBS) to measure DNA methylation in female guppies during distinct phases of infection by the guppy-specific *G. turnbulli*, and detected both a general epigenetic response across infection phases, and unique epigenetic responses in each phase. We found increased genomic DNA methylation levels in *infected* guppies, which is consistent with recent reports of genome-wide DNA methylation variation in infected mammal cells, in which increased DNA methylation was induced by parasite infection (Paschos & Allday 2010; Hari Dass & Vyas 2014). We also observed mean methylation differences of ~30% per DMR in skin tissue of *infected* versus *control* groups, a level that is on par with methylation responses previously observed for important phenotypic changes such as migration-related phenotypes (migratory vs. nonmigratory) in rainbow trout (Baerwald *et al.* 2015), and distinct caste phenotypes in honeybees (Herb *et al.* 2012). In addition to the moderate overall magnitude of methylation differences, we identified a number of DMRs that mapped to gene regulatory regions (CpG islands), and genomic regions close to immune genes. Hierarchical clustering based on methylation patterns alone was sufficient to differentiate between *infected* and *control* guppies, although there were individual exceptions (see below). Overall, our study provides the first investigation of epigenetic changes across distinct phases of infection by an ectoparasite, and identified DMRs that may be relevant to guppy immune response.

Linking differential DNA methylation to genomic architecture

We found approximately 50% of DMRs within or proximal to CpG islands, strongly suggesting that epigenetic modifications on the skin of *infected* guppies are influencing the transcriptional activity of associated genes. CpG islands at promoters remain predominantly unmethylated in somatic cells, and play a role in regulating transcription initiation in vertebrates, where approximately 70% of all annotated promoters are associated with CpG islands (Saxonov *et al.* 2006). Hypermethylation of these sites is typically associated with gene repression (Jones 2012). Thus, the general hypermethylation response observed is somewhat unexpected; we predicted increased gene expression would occur as the host mounts an immune response to infection, or during host angiogenesis and repair in response to parasite foraging on skin. This pattern highlights the importance of understanding the functional roles of loci overlapping with DMRs. For instance, in contrast to the predominant hypermethylation occurring in *infected* guppies relative to controls, it is notable that we observed a 40% decrease in DNA methylation in the *tenascin XB* promoter of infected fish

during the *recovery* phase compared to their controls. The *tenascin XB* locus is involved in wound healing and maintaining skin tissue structure (Gbadegesin *et al.* 2013), and hypomethylation of its promoter suggests that this gene is more stably expressed in the skin tissue of *infected* fish, and is thus likely to be involved in healing and skin repair after *Gyrodactylus* infection.

While DMCs were not significantly enriched in gene bodies, over half of all DMRs in our study were found to be overlapping with gene bodies, and all DMRs overlapping with immune gene locations were located within gene bodies. This is unexpected because, in contrast to promoters, gene bodies are typically CpG-poor, extensively methylated, and contain repetitive and transposable elements (Jones 2012). Although the functions of methylation in gene bodies remain unknown, there is some evidence suggesting that variation in DNA methylation in gene bodies can result in alternative splicing (Jones 2012), facilitate mutation by providing genetic variation during somatic hypermutation in immune genes (Racanelli & Rehermann 2006), and even regulate the activation of transposable elements to facilitate systemic responses to parasite infection (Wenzel & Piertney 2014). Indeed, when compared with controls, guppies in the *infected* group – regardless of infection phase – showed 43% higher DNA methylation levels in the gene body of the *protein phosphatase 2 regulatory subunit B'delta* (*PPP2R5D*). This gene controls substrate specificity and cellular localization, and plays a role as a regulator of tumorigenesis, drug resistance, and immune surveillance (Ruvolo 2016). Thus, methylation variation in this gene may result in a different transcript that facilitates guppies' resistance to infection, or over the long term, may produce new mutations for parasite-mediated selection to act on.

Immunological relevance of DMRs

Our gene ontology analysis identified multiple genes relevant to immune response within DMRs (Table 2.1). DMRs mapped to genes involved in antiviral response (Robertsen 2006), immune complexes (Schraml *et al.* 2006), antimicrobial peptide production (Fernandes & Smith 2004), and T-regulatory cell activation, central and peripheral tolerance establishment (Dougall *et al.* 1999; Theill *et al.* 2002; González-Suárez & Sanz-Moreno 2016). DMRs also mapped to genes that are mainly associated with the development and differentiation of leukocyte cells in the epithelium (e.g., mast cell, macrophage, neutrophil granulocyte), which is important for immune response because altering leukocyte cells can make consuming mucus energetically unfavourable for parasites (Jones 2001; Buchmann & Lindenstrøm 2002;

Dalgaard *et al.* 2003). The involvement of this collection of genes suggests that *Gyrodactylus* can induce diverse immune responses in guppies. Interestingly, only a few of the immune genes and GO terms were shared between individual phases of infection, which suggests that interactions between guppies and *Gyrodactylus* change throughout the development of infection. This observation is consistent with previous studies showing temporal shifts in gene expression in animals infected with intracellular parasites (Westermann *et al.* 2012; Choi *et al.* 2014). However, none of the immune-related DMRs overlapped with the promoters of immune genes, which does not provide evidence in support of our hypothesis that an active response of the host to infection would be reflected by hypomethylation in the promoters of immune genes.

Previous studies have identified several genes (e.g., MHC) that are under selection during *Gyrodactylus* infection (van Oosterhout *et al.* 2006; Fraser & Neff 2009; Fraser *et al.* 2010; Tonteri *et al.* 2010; Kjaerner-Semb *et al.* 2016), and that are differentially expressed in fish with different parasite loads (Lindenstrøm *et al.* 2004; Tadiso *et al.* 2011). Surprisingly, however, we found no overlap between these previously identified candidate genes for infection response and those mapping to DMRs in our study. This may suggest that the observed epigenetic differences are due to *trans*-acting genetic variants as opposed to more local *cis*-acting variants. However, the lack of overlap could also suggest that genetic and epigenetic variation represent independent mechanisms for facilitating adaptation or acclimation to infection (Klironomos *et al.* 2013). This possibility highlights the importance of studies such as ours that can uncover complementary sources of candidate loci relevant to immune response to parasite infections.

Functional associations between DMRs and immune response

Our KEGG analysis on annotated DMRs revealed important functional associations in canonical pathways and common gene networks. In particular, we identified several molecular pathways associated with disease and xenobiotic metabolism, including the following top three pathways that were shared by all infection phases: purine metabolism, aminobenzoate degradation, and thiamine metabolism (Table S2.3-S2.5, Supporting information). Purine metabolism can affect immunity, stress tolerance and resistance to infectious diseases in fish (Gil 2002; Li & Gatlin 2006; Dawood *et al.* 2017), and aminobenzoate degradation is associated with disease severity and stress response (Gevers *et al.* 2014). The thiamine metabolism pathway also helps to regulate the immune system

through the activation of immune cells (Manzetti *et al.* 2014). However, in general, the physiological processes revealed by our ontology and pathway analyses were diverse, and genes annotated with immune response were only a small subset of all genes that were differentially methylated. This is consistent with the view that parasite infection impacts physiological condition through a wide range of cellular processes in addition to strict immune responses (Hill 2011; Wenzel & Piertney 2014).

Unique epigenetic responses in distinct infection phases

Among the different infection phases, we found that the number of DMRs was highest in the *recovery* phase. This might be expected, as higher gene regulation during this phase could represent a fully mounted and active response to *Gyrodactylus,* and would also explain the reduction in infection levels observed in these fish. We also found that the majority of DMRs were unique to only a single infection phase. This may reflect changes in methylation driven by the dynamics of the infection cycle (e.g., Tadiso *et al.* 2011). However, it is also possible that some aspects of infection dynamics are produced by changes in methylation. Indeed, we found some DMRs that mapped to genes that are themselves involved in regulating epigenetic modifications (e.g., N-lysine methyltransferase), and active methylation changes at these epiloci may regulate methylation levels elsewhere in the genome. There is mounting evidence suggesting that infection-associated changes in methylation patterns are not primarily driven by the host, but are rather adaptive parasite-induced manipulations (Schmid Hempel & Schmid-Hempel 2011). Nonetheless, our finding that differential methylation was greatest in fish that were able to limit and reduce infection could suggest that hosts are also able to induce their own adaptive methylation responses.

Potential caveats

Our experiment has some limitations that should be noted. The most common technique used to measure DNA methylation is methylation-sensitive-AFLP (Schrey *et al.* 2013), which identifies global methylation changes, but does not provide the single-base resolution needed to extract functional genomic information. The RRBS approach used in our study has several advantages compared to the MS-AFLP, with the most important being its single-nucleotide resolution and greater genomic coverage, which allows for a more complete analysis of genomic sequences underlying differentially methylated regions. However, the RRBS approach does require a well-assembled reference genome with good annotations for

alignment and functional analysis purposes. Because the guppy genome was annotated using the 'The NCBI Eukaryotic Genome Annotation Pipeline' with predicted functions but no gene ontology terms (Künstner *et al.* 2016), we could not always convert gene IDs from the guppy genome to ENSEMBL IDs in corresponding model species, resulting in an incomplete analysis of DMR-associated gene functions.

Second, although we measured parasite load, we did not measure the infection-induced phenotypic change of each fish, for example, mate choice, and our sampling procedure is terminal, so the direct associations between epigenetic and phenotypic variation and fitness remain unclear. The relationships between the hyper/hypomethylation of DMRs and gene expression of the loci that they map to are complex, and although we have speculated about the potential effects of DMRs on phenotype (and possibly fitness) based on genomic architecture and gene ontology, these inferences are necessarily speculative pending future work to directly investigate the phenotypic and fitness consequences of differential methylation.

Third, although our analyses only included reads that uniquely aligned to one location, we cannot distinguish all gene family members due to the extensive repetitive sequences (approximately 20%) in the guppy genome (Künstner *et al.* 2016). Several of the DMR-associated genes we observed are members of well-known gene families (e.g., *WNT2*, *ANXA5*, *CDCA8*). Thus, it is possible that in some cases we have misidentified the specific gene family member showing differential methylation. In the absence of experiments (e.g., bisulphite cloning) to distinguish gene members, results should be interpreted with caution.

Fourth, we only used skin tissue from adult individuals, whereas DNA methylation is known to be specific to tissue and development stage, and infection can induce both local and systematic immune response. Thus, using different tissues at the same infection phase, or sampling fish from other developmental stages, might have allowed us a more complete understanding of immune responses.

In summary, although we provided evidence for broad epigenetic changes that were induced by parasite infection, and differed across infection phases, the evolutionary consequences of these changes remain unexplored. The relevance of epigenetic variation for evolution rests on whether epigenetically induced responses are under genetic control, and whether these responses can improve fitness (Richards 2006; Hu & Barrett 2017). It has been suggested that certain epigenetically induced responses in animals can be inherited over several generations in the laboratory (Daxinger & Whitelaw 2012; Lim & Brunet 2013; Heard & Martienssen 2014); however, the stability of these responses over longer

evolutionary timescales is unclear. If the methylation patterns observed here are heritable across generations, they could potentially increase fitness by providing offspring with innate resistance to parasites.

Conclusions

Consistent with previous studies of endoparasites, here we present the first evidence that ectoparasites can have important effects on genomic DNA methylation of their host. Our genome-wide methylation data shows significant epigenetic changes in guppies infected with an isogenic strain of the guppy-specific *G. turnbulli,* and indicates that these changes vary across different phases of infection. We found an underrepresentation of methylation variation in promoters, and over half of the DMRs overlapping with gene bodies, suggesting an important role of gene body methylation in host-parasite interactions. While discriminating between the causes and consequences of methylation variation is challenging, the high number of DMRs in fish showing successful recovery from infection suggests that these modifications could potentially be driven by an active response of the host as opposed to being regulated by the parasite. Our study adds to the large body of literature on guppy-*Gyrodactylus* interactions by characterising the epigenetic modifications associated with infection dynamics, and demonstrates that epigenetic modifications in guppies play an important role in the immune response to *Gyrodactylus*. Further investigation of DNA methylation patterns across natural host and parasite populations may be key to explaining the variation in resistance to infection observed in nature, as well as the evolution of complex phenotypic traits in the context of host-parasite interactions.

Acknowledgements

The authors would like to thank members of DryBar, Amanda Lea, Daniel Schoen, and the three anonymous reviewers for useful feedback on the manuscript. This work was supported by China Scholarship Council Fellowship 201406350023 to JH, a Postdoctoral Fellowship to FPJ and a Doctoral Award to LB by the NSERC CREATE 2015-466283, and a NSERC Discovery Grant and Canada Research Chair to RDHB.

Data Accessibility

Raw Illumina sequencing reads for the 17 analysed individuals can be downloaded from the NCBI Short Read Archive (SRA accession: SRP145142).

Author Contributions

JH and RDHB conceived the study. JH, FPJ, and LB ran the experiment and collected the data, JH analysed the data, and JH wrote the manuscript with input from FPJ, LB, and RDHB. The authors have no conflicts of interest.

Table 2.1 Gene Ontology (GO) terms in immune-related genes

Transcript ID †	Transcript name	GO ID	GO term	Meth Diff ‡
Pooling infected vs. controls				
XM_008419655.2	PREDICTED: *Poecilia reticulata interferon gamma receptor 1* (*ifngr1*), mRNA	GO:0030097	Hemopoiesis	-24.5
Early infection				
XM_008405926.2	PREDICTED: *Poecilia reticulata interferon regulatory factor 8* (*irf8*), mRNA	GO:0014005	Microglia development	-28.4
		GO:0045649	Regulation of macrophage differentiation	
		GO:0045658	Regulation of neutrophil differentiation	
XM_008416346.2	PREDICTED: *Poecilia reticulata mast/stem cell growth factor receptor kita-like* (LOC103468941), transcript variant X1, mRNA	GO:0038093	Fc receptor signaling pathway	-27.6
XM_008426764.1	PREDICTED: *Poecilia reticulata ribosomal protein L35a* (*rpl35a*), mRNA	GO:0030218	Erythrocyte differentiation	29.2
XM_008431017.1	PREDICTED: *Poecilia reticulata protein C-ets-1-like* (*ets1*), transcript variant X1, mRNA	GO:0030223	Neutrophil differentiation	23.1
		GO:0060217	Hemangioblast cell differentiation	
XM_008432066.1	PREDICTED: *Poecilia reticulata cholecystokinin-like* (LOC103478316), mRNA	GO:0006955	Immune response	-20.1
XM_008432236.2	PREDICTED: *Poecilia reticulata N-*	GO:0045576	Mast cell activation	23.1

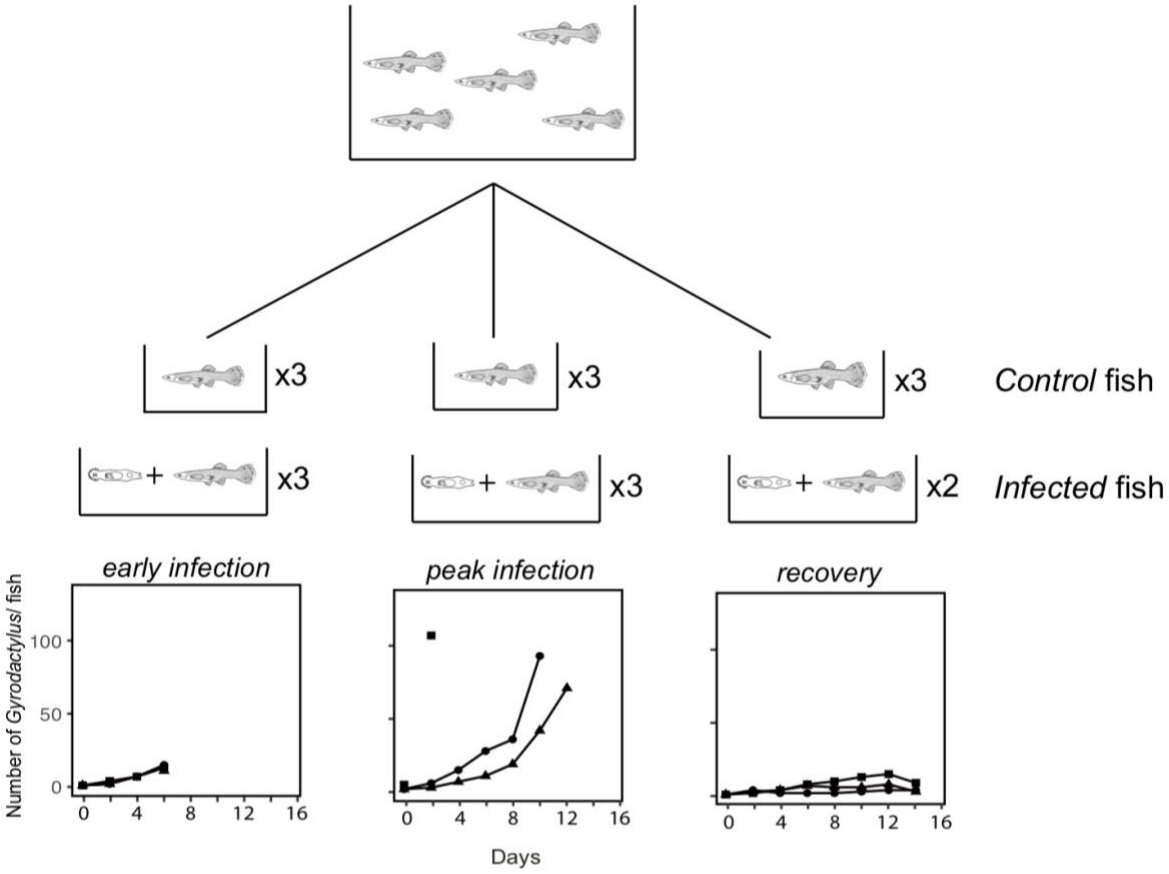
	myc downstream regulated 1 (*ndrg1*), mRNA			
XM_017310387.1	PREDICTED: *Poecilia reticulata L-amino-acid oxidase-like* (LOC103480925), transcript variant X2, mRNA	GO:0045087	Innate immune response	-23.0
Peak infection				
XM_008398390.1	PREDICTED: *Poecilia reticulata ribosomal protein S29* (*rps29*), mRNA	GO:0048821 GO:0060218	Erythrocyte development Hematopoietic stem cell differentiation	-25.7
XM_008419655.2	PREDICTED: *Poecilia reticulata interferon gamma receptor 1* (*ifngr1*), mRNA	GO:0030097	Hemopoiesis	-23.6
XM_008432236.2	PREDICTED: *Poecilia reticulata N-myc downstream regulated 1* (*ndrg1*), mRNA	GO:0045576	Mast cell activation	24.4
XM_008435495.2	PREDICTED: *Poecilia reticulata tumour necrosis factor superfamily member 12* (*tnfsf12*), mRNA	GO:0006955	Immune response	-20.8
XM_008437577.2	PREDICTED: *Poecilia reticulata tumour necrosis factor superfamily member 11* (*tnfsf11*), mRNA	GO:0006955	Immune response	-31.6
Recovery				
XM_008398848.1	PREDICTED: *Poecilia reticulata ubiquitin conjugating enzyme E2 D3* (*ube2d3*), transcript variant X1, mRNA	GO:0002223 GO:0035666 GO:0038095	Stimulatory C-type lectin receptor signaling pathway TRIF-dependent toll-like receptor signaling pathway Fc-epsilon receptor signaling pathway	21.5

		GO:0050852	T cell receptor signaling pathway	
XM_008409600.2	PREDICTED: *Poecilia reticulata GATA-binding factor 2-like* (LOC103465060), *transcript variant X3*, mRNA	GO:0048821 GO:0060215	Erythrocyte development Primitive hemopoiesis	32.3
XM_008419655.2	PREDICTED: *Poecilia reticulata interferon gamma receptor 1* (*ifngr1*), mRNA	GO:0030097	Hemopoiesis	-41.8
XM_008421008.2	PREDICTED: *Poecilia reticulata macrophage erythroblast attacher* (maea), mRNA	GO:0043249	Erythrocyte maturation	-21.2
XM_008426764.1	PREDICTED: *Poecilia reticulata ribosomal protein L35a* (*rpl35a*), mRNA	GO:0030218	Erythrocyte differentiation	27.2
XM_008427554.2	*Poecilia reticulata junctional adhesion molecule 2* (*jam2*), transcript variant X2, mRNA	GO:0048534	Hematopoietic or lymphoid organ development	30.5
XM_017309240.1	PREDICTED: *Poecilia reticulata protein tyrosine phosphatase, non-receptor type 6* (*ptpn6*), transcript variant X3, mRNA	GO:0045087	Innate immune response	-37.5

† DMRs overlapping with genes, or regions that are 5 kb up- or downstream of gene locations are labelled in bold.

‡ Per cent methylation differences averaged from all CpG sites within the defined region. The comparison is between *infected* and *control* fish with positive values representing increased methylation for *infected* fish, and negative values representing decreased methylation for *infected* fish.

a


x3
x3
x3
Control fish
x3
x3
x2
Infected fish
early infection
peak infection
recovery
Number of Gyrodactylus/ fish
100
50
0
0
4
8
12
16
Days

b


4 Ultraviolet Sterilizer
1 Particulate Filter Pad
3 Carbon Filters
2 Fluidized Bed Biological Filter

Fig. 2.1 Overview of the experimental design used to set up the infection trial. (a) Guppies from a genetically homogeneous population were randomly assigned to *control* or *infected* groups in three infection phases: *early infection*, *peak infection,* and *recovery* based on the number of parasites on each fish. Each *control* or *infected* group contains three replicate fish (*control* group in *recovery* contains two fish), and individual fish were kept in separated tanks until the end of experiment. (b) The flow through system (Aquaneering Inc., San Diego, CA, USA. Image source: http://www.aquaneering.com/zebrafish_stand_alone_systems.php). Water is filtered by particulate, biological, and carbon filters, and flows into individual tanks. The in-line UV steriliser lamp provides UV light preventing potential movement of parasites between tanks.

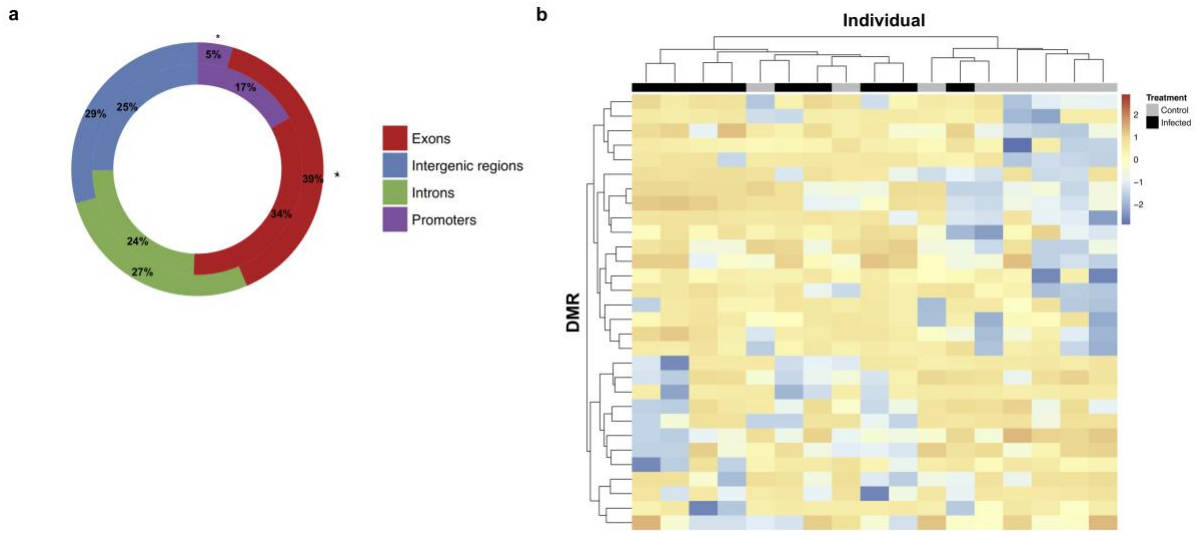


Fig. 2.2 (a) The proportion of genomic features (promoters, exons, introns, or intergenic regions) in DMCs compared with the genomic features of all filtered CpG sites. The outer ring describes the locations of DMCs, the inner ring describes the features of filtered CpG sites. Asterisks denote significant differences between the features of DMCs versus the features of filtered CpG sites using a *G*-test at $P < 0.05$. (b) Heatmap of methylation levels of the 30 DMRs when comparing *infected* versus *control* fish pooling across all phases. Each column represents a colour-coded individual: black for *infected* fish, and grey for *control* fish. Each row represents one of the DMRs, which are clustered based on the similarities of the methylation patterns between individuals. Darker red indicates greater methylation in an individual for that DMR. Darker blue indicates lesser methylation in an individual for that DMR. Individual dendrogram positions are based on overall methylation patterns across the 30 DMRs.

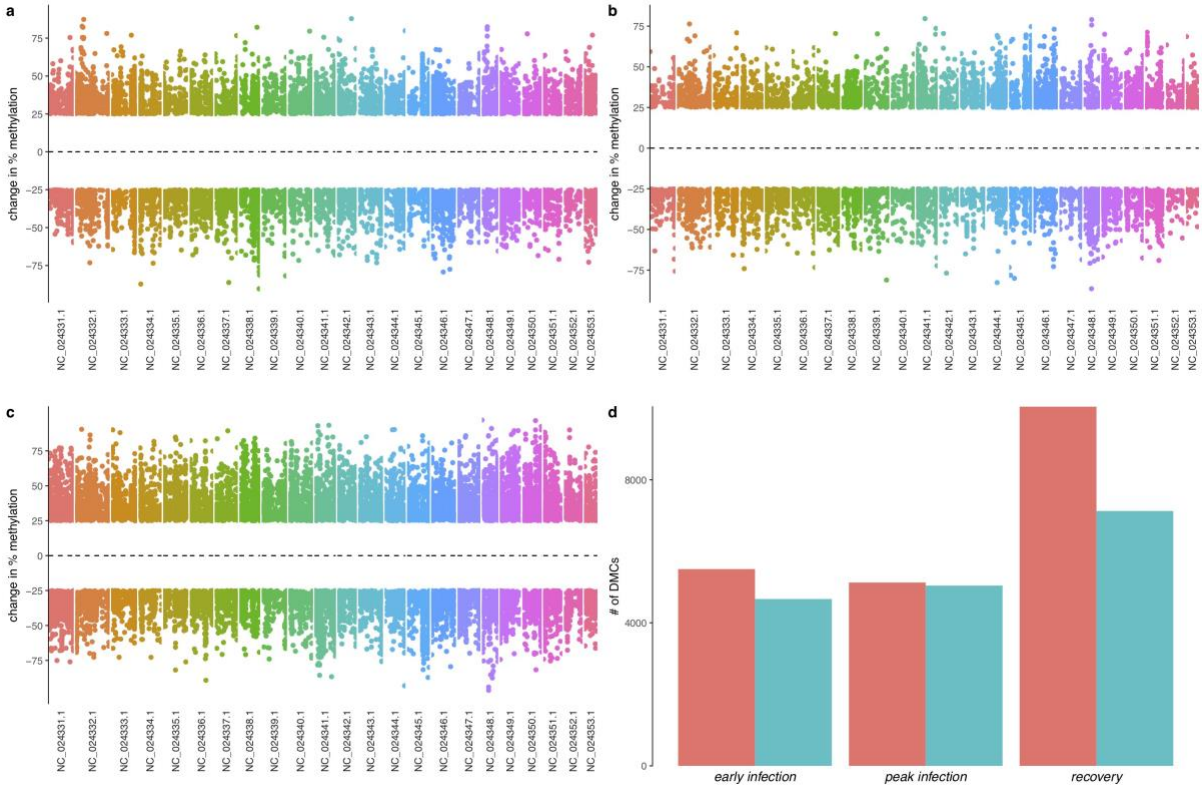


Fig. 2.3 (a–c) Manhattan plots of the chromosomal positions of methylated CpG loci that differed significantly between *infected* and *control* fish in (a) *early infection*, (b) *peak infection* or (c) *recovery*. Each point represents a single DMC. The y-axis presents the difference in percentage methylation for that DMC in *infected* fish relative to the *control* fish. Only DMCs with more than 25% change in methylation are shown. Points above and below the horizontal dashed line are hypermethylated and hypomethylated loci, respectively. (d) Number of hypermethylated (red) and hypomethylated (blue) DMCs across the guppy genome in *early infection*, *peak infection*, or *recovery* phase.

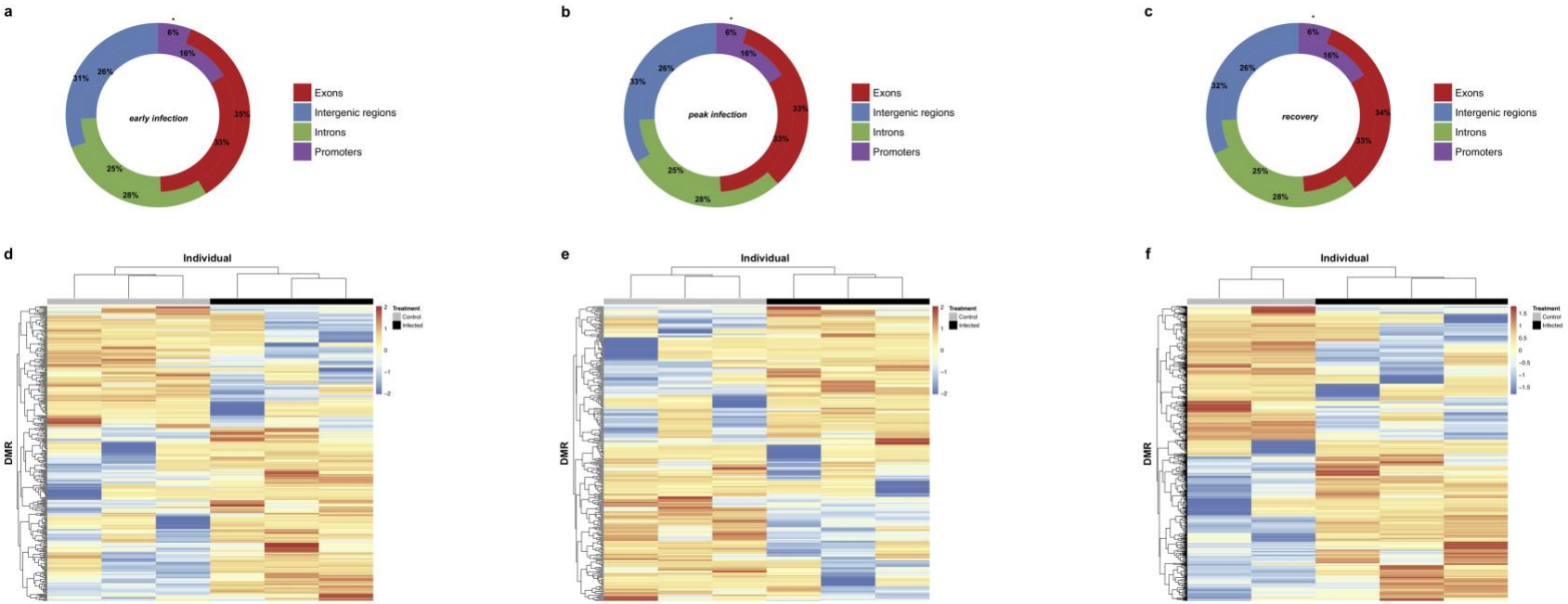


Fig. 2.4 (a-c) The proportion of genomic features (promoters, exons, introns, or intergenic regions) in DMCs compared with genomic features of filtered CpG sites in (a) *early infection*, (b) *peak infection*, and (c) *recovery*. Outer rings describe the locations of DMCs, inner rings describe the features of all filtered CpG sites. Asterisks denote significant differences between the features of DMCs versus the features of filtered CpG sites using a *G*-test at $P < 0.01$. (d-f) Heatmap of methylation levels of DMRs when comparing *infected* versus *control* fish in (d) *early infection*, (e) *peak infection*, and (f) *recovery*. Each column represents a colour-coded individual: black for *infected* fish, and grey for *control* fish. Each row represents one of the DMRs identified within a phase. DMRs are clustered based on the similarities of the methylation patterns between individuals. The darker the red, the more methylated that individual is for that DMR. The darker the blue, the less methylated that individual is for that DMR. Individual dendrogram positions are based on their overall methylation patterns across DMRs identified within a phase.

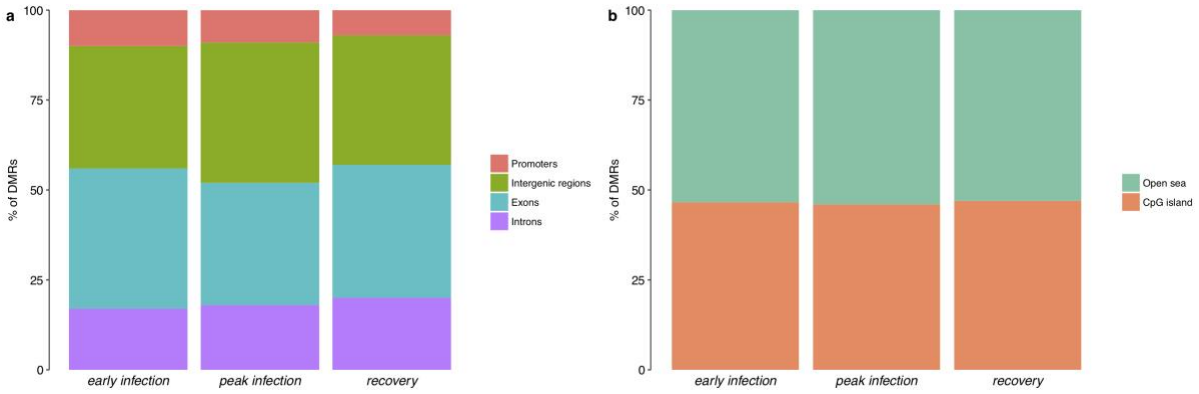


Fig. 2.5 Genomic architecture of DMRs. (a) Proportion of DMRs overlapping with genomic features (promoters, exons, introns, or intergenic regions) in *early infection*, *peak infection*, and *recovery* phases. Overlapping genomic features were given the precedence promoters > exons > introns > intergenic regions. (b) Proportion of DMRs that are within or proximal to CpG islands or open sea in *early infection*, *peak infection*, and *recovery* phases.

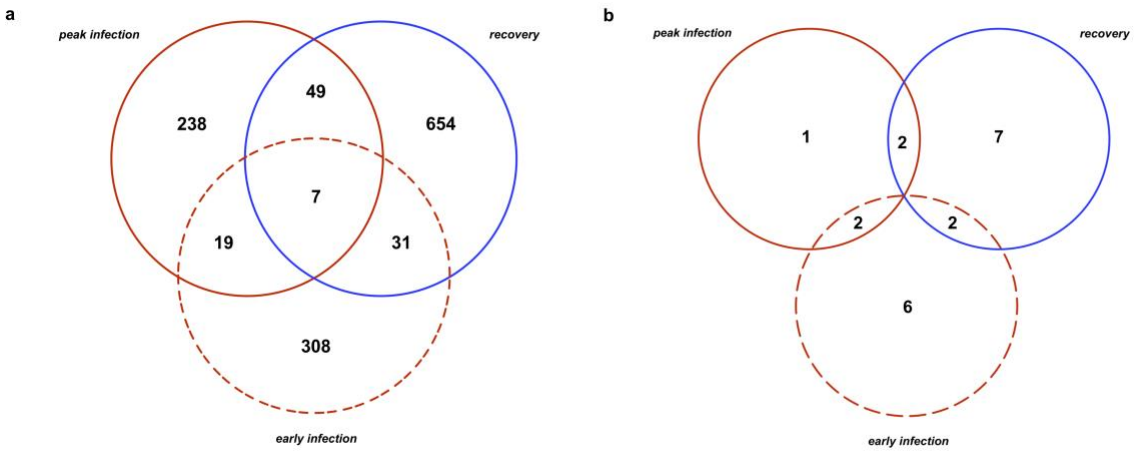


Fig. 2.6 Venn diagrams showing characteristics of DMRs. (a) The number of DMRs that are unique in each phase, and shared between different phases. (b) The number of GO terms annotated by immune-related DMRs that are unique in each phase, and shared between different phases.

References

Akalin A, Franke V, Vlahoviček K, Mason CE, Schübeler D (2015) genomation: a toolkit to summarize, annotate and visualize genomic intervals. *Bioinformatics*, **31**, 1127-1129.

Akalin A, Kormaksson M, Li S *et al.* (2012) methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology*, **13**, R87.

Ángeles Esteban M (2012) An overview of the immunological defenses in fish skin. *ISRN Immunology*, **2012**, 1-29.

Baerwald MR, Meek MH, Stephens MR *et al.* (2015) Migration-related phenotypic divergence is associated with epigenetic modifications in rainbow trout. *Molecular Ecology*, **25**, 1785-1800.

Bakke TA, Cable J, Harris PD (2007) The Biology of Gyrodactylid monogeneans: the "Russian-doll killers". *Advances in Parasitology*, **64**, 161-460.

Boyko A, Kovalchuk I (2011) Genetic and epigenetic effects of plant-pathogen interactions: an evolutionary perspective. *Molecular plant*, **4**, 1014-1023.

Boyle P, Clement K, Gu H *et al.* (2012) Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. *Genome Biology*, **13**, R92.

Buchmann K, Bresciani J (1997) Microenvironment of *Gyrodactylus derjavini* on rainbow trout *Oncorhynchus mykiss*: association between mucous cell density in skin and site selection. *Parasitology Research*, **84**, 17-24.

Buchmann K, Lindenstrøm T (2002) Interactions between monogenean parasites and their fish hosts. *International Journal for Parasitology*, **32**, 309-319.

Cable J, van Oosterhout C (2007) The impact of parasites on the life history evolution of guppies (*Poecilia reticulata*): The effects of host size on parasite virulence. *International Journal for Parasitology*, **37**, 1449-1458.

Cheeseman K, Weitzman JB (2015) Host-parasite interactions: an intimate epigenetic relationship. *Cellular Microbiology*, **17**, 1121-1132.

Choi Y-J, Aliota MT, Mayhew GF, Erickson SM, Christensen BM (2014) Dual RNA-seq of parasite and host reveals gene expression dynamics during filarial worm-mosquito interactions. *PLoS Neglected Tropical Diseases*, **8**, e2905.

Conesa A, Gotz S, Garcia-Gomez JM *et al.* (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, **21**, 3674-3676.

Conrath U (2011) Molecular aspects of defence priming. *Trends in Plant Science*, **16**, 524-531.

Cusack R, Cone DK (1986) A review of parasites as vectors of viral and bacterial diseases of fish. *Journal of Fish Diseases*, **9**, 169-171.

Dalgaard MB, Nielsen CV, Buchmann K (2003) Comparative susceptibility of two races of *Salmo salar* (Baltic Lule river and Atlantic Conon river strains) to infection with *Gyrodactylus salaris*. *Diseases of Aquatic Organisms*, **53**, 173-176.

Dargent F, Scott ME, Hendry AP, Fussmann GF (2013) Experimental elimination of parasites in nature leads to the evolution of increased resistance in hosts. *Proceedings of the Royal Society of London Series B*, **280**, 20132371.

Dawood MAO, Koshio S, Esteban MÁ (2017) Beneficial roles of feed additives as immunostimulants in aquaculture: a review. *Reviews in Aquaculture*, **0**, 1-25.

Daxinger L, Whitelaw E (2012) Understanding transgenerational epigenetic inheritance via the gametes in mammals. *Nature Reviews Genetics*, **13**, 153-162.

Dougall WC, Glaccum M, Charrier K *et al.* (1999) RANK is essential for osteoclast and lymph node development. *Genes & Development*, **13**, 2412-2424.

Dowen RH, Pelizzola M, Schmitz RJ *et al.* (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.

Endler JA (1995) Multiple-trait coevolution and environmental gradients in guppies. *Trends in Ecology & Evolution*, **10**, 22-29.

Fernandes JMO, Smith VJ (2004) Partial purification of antibacterial proteinaceous factors from erythrocytes of Oncorhynchus mykiss. *Fish & Shellfish Immunology*, **16**, 1-9.

Fraser BA, Neff BD (2009) MHC class IIB additive and non-additive effects on fitness measures in the guppy Poecilia reticulata. *Journal of Fish Biology*, **75**, 2299-2312.

Fraser BA, Ramnarine IW, Neff BD (2010) Selection at the MHC class IIB locus across guppy (*Poecilia reticulata*) populations. *Heredity*, **104**, 155-167.

Gbadegesin RA, Brophy PD, Adeyemo A *et al.* (2013) *TNXB* mutations can cause vesicoureteral reflux. *Journal of the American Society of Nephrology*, **24**, 1313-1322.

Gevers D, Kugathasan S, Denson Lee A *et al.* (2014) The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host & Microbe*, **15**, 382-392.

Gheorghiu C, Marcogliese DJ, Scott ME (2012) Waterborne zinc alters temporal dynamics of guppy *Poecilia reticulata* epidermal response to *Gyrodactylus turnbulli* (Monogenea). *Diseases of Aquatic Organisms*, **98**, 143-153.

Gil A (2002) Modulation of the immune response mediated by dietary nucleotides. *European Journal of Clinical Nutrition*, **56**, S1.

Gómez-Díaz E, Jorda M, Peinado MA, Rivero A (2012) Epigenetics of host-pathogen interactions: the road ahead and the road behind. *PLoS Pathogens*, **8**, e1003007.

González-Suárez E, Sanz-Moreno A (2016) RANK as a therapeutic target in cancer. *The FEBS Journal*, **283**, 2018-2033.

Gotanda KM, Delaire LC, Raeymaekers JA *et al.* (2013) Adding parasites to the guppy-predation story: insights from field surveys. *Oecologia*, **172**, 155-166.

Götz S, García-Gómez JM, Terol J *et al.* (2008) High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research*, **36**, 3420-3435.

Gu H, Smith ZD, Bock C *et al.* (2011) Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nature Protocols*, **6**, 468-481.

Hamilton WD (1980) Sex versus non-sex versus parasite. *Oikos*, **35**, 282-290.

Hari Dass SA, Vyas A (2014) *Toxoplasma gondii* infection reduces predator aversion in rats through epigenetic modulation in the host medial amygdala. *Molecular Ecology*, **23**, 6114-6122.

Hatcher MJ, Dick JTA, Dunn AM (2006) How parasites affect interactions between competitors and predators. *Ecology Letters*, **9**, 1253-1271.

Heard E, Martienssen RA (2014) Transgenerational epigenetic inheritance: myths and mechanisms. *Cell*, **157**, 95-109.

Herb BR, Wolschin F, Hansen KD *et al.* (2012) Reversible switching between epigenetic states in honeybee behavioral subcastes. *Nature Neuroscience*, **15**, 1371-1373.

Hill GE (2011) Condition-dependent traits as signals of the functionality of vital cellular processes. *Ecology Letters*, **14**, 625-634.

Holeski LM, Jander G, Agrawal AA (2012) Transgenerational defense induction and epigenetic inheritance in plants. *Trends in Ecology & Evolution*, **27**, 618-626.

Houde AE (1997) *Sex, color, and mate choice in guppies.* Princeton University Press, Princeton, NJ.

Hu J, Barrett RDH (2017) Epigenetics in natural animal populations. *Journal of Evolutionary Biology*, **30**, 1612-1632.

Jones PA (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*, **13**, 484-492.

Jones SRM (2001) The occurrence and mechanisms of innate immunity against parasites in fish. *Developmental & Comparative Immunology*, **25**, 841-852.

Kania P, Larsen TB, Ingerslev HC, Buchmann K (2007) Baltic salmon activates immune relevant genes in fin tissue when responding to *Gyrodactylus salaris* infection. *Diseases of Aquatic Organisms*, **76**, 81-85.

Kennedy CEJ, Endler JA, Poynton SL, McMinn H (1987) Parasite load predicts mate choice in guppies. *Behavioral Ecology and Sociobiology*, **21**, 291-295.

Kjaerner-Semb E, Ayllon F, Furmanek T *et al.* (2016) Atlantic salmon populations reveal adaptive divergence of immune related genes - A duplicated genome under selection. *BMC Genomics*, **17**, 610.

Klironomos FD, Berg J, Collins S (2013) How epigenetic mutations can affect genetic evolution: model and mechanism. *BioEssays*, **35**, 571-578.

Kolluru GR, Grether GF, South SH *et al.* (2006) The effects of carotenoid and food availability on resistance to a naturally occurring parasite (*Gyrodactylus turnbulli*) in guppies (*Poecilia reticulata*). *Biological Journal of the Linnean Society*, **89**, 301-309.

Kotob MH, Menanteau-Ledouble S, Kumar G, Abdelzaher M, El-Matbouli M (2016) The impact of co-infections on fish: a review. *Veterinary Research*, **47**, 98.

Krueger F, Andrews S (2011) Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics*, **27**, 1571-1572.

Kumar S, Raman RP, Prasad KP *et al.* (2017) Modulation of innate immune responses and induction of oxidative stress biomarkers in *Pangasianodon hypophthalmus* following an experimental infection with dactylogyrid monogeneans. *Fish & Shellfish Immunology*, **63**, 334-343.

Künstner A, Hoffmann M, Fraser BA *et al.* (2016) The genome of the Trinidadian guppy, *Poecilia reticulata*, and variation in the Guanapo population. *PLoS ONE*, **11**, e0169087.

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods*, **9**, 357-359.

Lazzaro BP, Flores HA, Lorigan JG, Yourth CP (2008) Genotype-by-environment interactions and adaptation to local temperature affect immunity and fecundity in *Drosophila melanogaster*. *PLoS Pathogens*, **4**, e1000025.

Le Luyer J, Laporte M, Beacham TD *et al.* (2017) Parallel epigenetic modifications induced by hatchery rearing in a Pacific salmon. *Proceedings of the National Academy of Sciences of the United States of America*, **114**, 12964-12969.

Li P, Gatlin DM (2006) Nucleotide nutrition in fish: Current knowledge and future applications. *Aquaculture*, **251**, 141-152.

Li S, Garrett-Bakelman FE, Akalin A *et al.* (2013) An optimized algorithm for detecting and annotating regional differential methylation. *BMC Bioinformatics*, **14**, 1-9.

Lim JP, Brunet A (2013) Bridging the transgenerational gap with epigenetic memory. *Trends in Genetics*, **29**, 176-186.

Lindenstrøm T, Buchmann K, Secombes CJ (2003) *Gyrodactylus derjavini* infection elicits IL-1*β* expression in rainbow trout skin. *Fish & Shellfish Immunology*, **15**, 107-115.

Lindenstrøm T, Secombes CJ, Buchmann K (2004) Expression of immune response genes in rainbow trout skin induced by *Gyrodactylus derjavini* infections. *Veterinary Immunology and Immunopathology*, **97**, 137-148.

López S (1999) Parasitized female guppies do not prefer showy males. *Animal Behaviour*, **57**, 1129-1134.

Magurran AE (2005) *Evolutionary ecology: the Trinidadian guppy.* Oxford University Press, New York, NY.

Manzetti S, Zhang J, van der Spoel D (2014) Thiamin function, metabolism, uptake, and transport. *Biochemistry*, **53**, 821-835.

Marr AK, MacIsaac JL, Jiang R *et al.* (2014) *Leishmania donovani* infection causes distinct epigenetic DNA methylation changes in host macrophages. *PLoS Pathogens*, **10**, e1004419.

Matejusová I, Felix B, Sorsa-Leslie T *et al.* (2006) Gene expression profiles of some immune relevant genes from skin of susceptible and responding Atlantic salmon (*Salmo salar* L.) infected with *Gyrodactylus salaris* (Monogenea) revealed by suppressive subtractive hybridisation. *International Journal for Parasitology*, **36**, 1175-1183.

McCarthy NS, Melton PE, Cadby G *et al.* (2014) Meta-analysis of human methylation data for evidence of sex-specific autosomal patterns. *BMC Genomics*, **15**, 981.

Meissner A, Mikkelsen T, Gu H *et al.* (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*, **454**, 766 - 770.

Paschos K, Allday MJ (2010) Epigenetic reprogramming of host genes in viral and microbial pathogenesis. *Trends in Microbiology*, **18**, 439-447.

Paterson S, Piertney SB (2011) Frontiers in host-parasite ecology and evolution. *Molecular Ecology*, **20**, 869-871.

Penczykowski RM, Laine AL, Koskella B (2016) Understanding the ecology and evolution of host-parasite interactions across scales. *Evolutionary Applications*, **9**, 37-52.

Pérez-Jvostov F, Hendry AP, Fussmann GF, Scott ME (2012) Are host-parasite interactions influenced by adaptation to predators? A test with guppies and *Gyrodactylus* in experimental stream channels. *Oecologia*, **170**, 77-88.

Pérez-Jvostov F, Hendry AP, Fussmann GF, Scott ME (2017) Experimental assessment in nature of the ecological effects of a specialist parasite. *Copeia*, **105**, 494-503.

Poulin R, Thomas F (2008) Epigenetic effects of infection on the phenotype of host offspring: Parasites reaching across host generations. *Oikos*, **117**, 331-335.

Racanelli V, Rehermann B (2006) The liver as an immunological organ. *Hepatology*, **43**, S54-S62.

Reznick D, Endler JA (1982) The impact of predation on life history evolution in Trinidadian guppies (*Poecilia reticulata*). *Evolution*, **36**, 160-177.

Reznick DN, Shaw FH, Rodd FH, Shaw RG (1997) Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science*, **275**, 1934.

Richards EJ (2006) Inherited epigenetic variation — revisiting soft inheritance. *Nature Reviews Genetics*, **7**, 395-401.

Richards GR, Chubb JC (1996) Host response to initial and challenge infections, following treatment, of *Gyrodactylus bullatarudis* and *G. turnbulli* (Monogenea) on the guppy (*Poecilia reticulata*). *Parasitology Research*, **82**, 242-247.

Robert McMaster W, Morrison CJ, Kobor MS (2016) Epigenetics: A new model for intracellular parasite-host cell regulation. *Trends in Parasitology*, **32**, 515-521.

Robertsen B (2006) The interferon system of teleost fish. *Fish & Shellfish Immunology*, **20**, 172-191.

Rohland N, Reich D (2012) Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Research*, **22**, 939-946.

Ruvolo PP (2016) The broken "Off" switch in cancer signaling: PP2A as a regulator of tumorigenesis, drug resistance, and immune surveillance. *BBA Clinical*, **6**, 87-99.

Salinas I, Zhang Y-A, Sunyer JO (2011) Mucosal immunoglobulins and B cells of teleost fish. *Developmental & Comparative Immunology*, **35**, 1346-1365.

Saxonov S, Berg P, Brutlag DL (2006) A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proceedings of the National Academy of Sciences*, **103**, 1412-1417.

Schmid Hempel P, Schmid-Hempel P (2011) *Evolutionary parasitologythe integrated study of infections, immunology, ecology, and genetics.* Oxford University Press, New York, NY.

Schraml B, Baker MA, Reilly BD (2006) A complement receptor for opsonized immune complexes on erythrocytes from *Oncorhynchus mykiss* but not *Ictalarus punctatus*. *Molecular Immunology*, **43**, 1595-1603.

Schrey AW, Alvarez M, Foust CM *et al.* (2013) Ecological epigenetics: Beyond MS-AFLP. *Integrative and Comparative Biology*, **53**, 340-350.

Scott ME (1982) Reproductive potential of *Gyrodactylus bullatarudis* (Monogenea) on guppies (*Poecilia reticulata*). *Parasitology*, **85**, 217-236.

Scott ME (1985) Dynamics of challenge infections of *Gyrodactylus bullatarudis* Turnbull (Monogenea) on guppies, *Poecilia reticulata* (Peters). *Journal of Fish Diseases*, **8**, 495-503.

Sessions OM, Tan Y, Goh KC *et al.* (2013) Host cell transcriptome profile during wild-type and attenuated dengue virus infection. *PLoS Neglected Tropical Diseases*, **7**, e2107.

Sheldon BC, Verhulst S (1996) Ecological immunology: costly parasite defences and trade-offs in evolutionary ecology. *Trends in Ecology & Evolution*, **11**, 317-321.

Silmon de Monerri NC, Kim K (2014) Pathogens hijack the epigenome. *The American Journal of Pathology*, **184**, 897-911.

Suzuki MM, Bird A (2008) DNA methylation landscapes: provocative insights from epigenomics. *Nature Reviews Genetics*, **9**, 465-476.

Tadiri CP, Dargent F, Scott ME (2013) Relative host body condition and food availability influence epidemic dynamics: A *Poecilia reticulata*-*Gyrodactylus turnbulli* host-parasite model. *Parasitology*, **140**, 343-351.

Tadiri CP, Scott ME, Fussmann GF (2016) Impact of host sex and group composition on parasite dynamics in experimental populations. *Parasitology*, **143**, 523-531.

Tadiso TM, Krasnov A, Skugor S *et al.* (2011) Gene expression analyses of immune responses in Atlantic salmon during early stages of infection by salmon louse (*Lepeophtheirus salmonis*) revealed bi-phasic responses coinciding with the copepod-chalimus transition. *BMC Genomics*, **12**, 141-141.

Takai D, Jones PA (2002) Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 3740-3745.

Taudt A, Colome-Tatche M, Johannes F (2016) Genetic sources of population epigenomic variation. *Nature Reviews Genetics*, **17**, 319-332.

Theill LE, Boyle WJ, Penninger JM (2002) RANK-L and RANK: T cells, bone loss, and mammalian evolution. *Annual Review of Immunology*, **20**, 795-823.

Tonteri A, VasemÄGi A, Lumme J, Primmer CR (2010) Beyond MHC: signals of elevated selection pressure on Atlantic salmon (*Salmo salar*) immune-relevant loci. *Molecular Ecology*, **19**, 1273-1282.

van Oosterhout C, Harris PD, Cable J (2003) Marked variation in parasite resistance between two wild populations of the Trinidadian guppy, *Poecilia reticulata* (Pisces: Poeciliidae). *Biological Journal of the Linnean Society*, **79**, 645-651.

van Oosterhout C, Joyce DA, Cummings SM *et al.* (2006) Balancing selection, random genetic drift, and genetic variation at the major histocompatibility complex in two wild populations of guppies (*Poecilia reticulata*). *Evolution*, **60**, 2562-2574.

van Oosterhout C, Mohammed RS, Hansen H *et al.* (2007) Selection by parasites in spate conditions in wild Trinidadian guppies (*Poecilia reticulata*). *International Journal for Parasitology*, **37**, 805-812.

Walker DL, Bhagwate AV, Baheti S *et al.* (2015) DNA methylation profiling: comparison of genome-wide sequencing methods and the Infinium Human Methylation 450 Bead Chip. *Epigenomics*, **7**, 1287-1302.

Wan ZY, Xia JH, Lin G *et al.* (2016) Genome-wide methylation analysis identified sexually dimorphic methylated regions in hybrid tilapia. *Scientific Reports*, **6**, 35903.

Wells PR, Cone DK (1990) Experimental studies on the effect of *Gyrodactylus colemanensis* and *G. salmonis* (Monogenea) on density of mucous cells in the epidermis of fry of *Oncorhynchus mykiss*. *Journal of Fish Biology*, **37**, 599-603.

Wenzel MA, Piertney SB (2014) Fine-scale population epigenetic structure in relation to gastrointestinal parasite load in red grouse (*Lagopus lagopus scotica*). *Molecular Ecology*, **23**, 4256-4273.

Westermann AJ, Barquist L, Vogel J (2017) Resolving host-pathogen interactions by dual RNA-seq. *PLoS Pathogens*, **13**, e1006033.

Westermann AJ, Gorski SA, Vogel J (2012) Dual RNA-seq of pathogen and host. *Nature Reviews Microbiology*, **10**, 618-630.

Youngblood B, Davis CW, Ahmed R (2010) Making memories that last a lifetime: heritable functions of self-renewing memory CD8 T cells. *International Immunology*, **22**, 797-803.

Zhu LJ (2013) Integrative analysis of ChIP-Chip and ChIP-Seq dataset. In: *Tiling Arrays: Methods and Protocols* (eds. Lee T-L, Shui Luk AC), pp. 105-124. Humana Press, Totowa, NJ.

Zhu LJ, Gazin C, Lawson ND *et al.* (2010) ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics*, **11**, 237.

Table S2.1 Read counts and alignments to guppy genome for infected and control individuals

ID #	Sample Name	Treatment	Infection Phase	Unfiltered	Filtered	Uniquely Mapped	No Alignments	No. CpG after alignment
Sample 1	E-I-Rep1	Infected	EEG	23,118,342	22,866,401	12,591,241	7,971,007	66,957,458
Sample 2	E-I-Rep2	Infected	EEG	20,786,624	20,577,355	11,877,921	6,426,157	59,267,920
Sample 3	E-I-Rep3	Infected	EEG	19,411,061	19,240,938	11,574,906	5,718,083	57,848,206
Sample 4	E-C-Rep1	Control	EEG	12,411,217	12,243,862	7,156,256	3,817,686	35,764,717
Sample 5	E-C-Rep2	Control	EEG	20,161,293	19,949,279	11,766,325	6,177,949	59,360,749
Sample 6	E-C-Rep3	Control	EEG	15,016,895	14,875,643	8,527,238	4,860,430	42,466,404
Sample 7	P-I-Rep1	Infected	LEG	17,407,057	17,209,114	10,224,141	5,275,600	51,492,858
Sample 8	P-I-Rep2	Infected	LEG	28,311,603	27,996,974	16,833,983	8,323,614	84,736,344
Sample 9	P-I-Rep3	Infected	LEG	18,518,593	18,295,807	11,235,244	5,153,965	56,320,345
Sample 10	P-C-Rep1	Control	LEG	18,628,737	18,450,882	10,821,816	5,783,380	54,127,121
Sample 11	P-C-Rep2	Control	LEG	19,209,037	19,023,902	10,752,603	6,412,067	54,227,314
Sample 12	P-C-Rep3	Control	LEG	20,120,286	19,917,303	11,861,419	6,060,602	59,296,294
Sample 13	R-I-Rep1	Infected	RECOV	15,668,723	15,487,655	9,273,719	4,600,367	46,442,273
Sample 14	R-I-Rep2	Infected	RECOV	25,035,142	24,824,797	13,891,634	8,618,985	71,547,545
Sample 15	P-I-Rep3	Infected	RECOV	14,964,281	14,793,404	9,315,137	3,922,181	46,399,366
Sample 16	R-C-Rep1	Control	RECOV	25,273,006	25,018,112	14,126,196	8,568,205	71,642,621
Sample 17	R-C-Rep2	Control	RECOV	23,583,168	23,341,705	13,792,177	7,186,243	70,186,838

Table S2.2 Differentially methylated regions (DMRs) when pooling *infected* and *control* guppies across all phases

Symbol	Entrez gene name	Meth diff †	*P*-value	*Q*-value
IFNGR1	Interferon gamma receptor 1	-24.5	8.2E-19	1.2E-18
LOC103461906	Tripartite motif-containing protein 54-like	-28.5	7.3E-88	1.1E-86
WNT2	Wnt family member 2	-24.5	1.5E-33	3.4E-33
LOC103466185	Seipin-like	-22.0	3.7E-18	5.3E-18
HAL	Histidine ammonia-lyase	-24.5	9.1E-26	1.5E-25
XPC	XPC complex subunit, DNA damage recognition and repair factor	-21.3	1.6E-74	1.2E-73
LOC103469125	Cerebellar degeneration-related protein 2-like	21.6	4.5E-52	1.6E-51
LOC103471094	Glutamate receptor 3	23.7	2.8E-36	6.6E-36
LOC103471608	Zinc finger protein ZIC 4-like	-24.6	1.5E-31	3.2E-31
LOC103471718	Sodium-dependent phosphate transport protein 2A	-22.1	2.2E-06	2.4E-06
ANXA5	Annexin A5	29.8	6.1E-54	2.5E-53
CDCA8	Cell division cycle associated 8	-24.8	9.2E-81	1.1E-79
PID1	Phosphotyrosine interaction domain containing 1	24.2	1.9E-41	5.1E-41
PID1	Phosphotyrosine interaction domain containing 1	25.3	4.5E-38	1.1E-37
LOC103474836	Complement C1q-like protein 2	27.2	7.4E-29	1.4E-28
LOC103475725	Protein phosphatase 1B-like	22.5	3.5E-29	7.0E-29
LOC103475725	Protein phosphatase 1B-like	33.3	3.8E-22	6.0E-22
LOC103475842	Voltage-gated potassium channel subunit beta-2-like	28.9	1.8E-16	2.5E-16
PPP2R5D	Protein phosphatase 2 regulatory subunit B'delta	44.8	4.7E-106	1.2E-104
LOC103477958	Zinc-binding protein A33-like	28.8	2.0E-59	9.1E-59
LOC103478552	Rap guanine nucleotide exchange factor 5-like	31.0	9.9E-67	5.6E-66
LOC103478903	Leucine-rich repeat-containing protein 24-like	25.6	1.1E-10	1.3E-10
LOC103480967	Complement factor I	30.8	6.9E-76	5.6E-75
LOC103481898	cGMP-dependent protein kinase 1-like	39.6	6.5E-68	3.9E-67
HADHB	Hydroxyacyl-CoA dehydrogenase/3-ketoacyl-CoA thiolase/enoyl-CoA hydratase (trifunctional protein), beta subunit	24.6	3.4E-32	7.5E-32
LOC103457724	E3 ubiquitin-protein ligase rnf146-like	26.1	4.2E-90	7.7E-89
ATAD2B	ATPase family, AAA domain containing 2B	-23.3	4.2E-44	1.2E-43
—	Unknown	-21.8	2.8E-51	9.7E-51

—	Unknown	-20.4	4.0E-39	1.0E-38
—	Unknown	22.4	1.1E-63	5.4E-63

[†]Per cent methylation differences averaged from all CpG sites within the defined region. The comparison is between *infected* and *control* fish with positive values representing increased methylation in *infected* fish and negative values representing decreased methylation in *infected* fish.

Table S2.3 Canonical pathways identified via KEGG in the *early infection* phase

Canonical pathway name	No. Sequences	No. Enzymes
Purine metabolism	7	4
Thiamine metabolism	6	1
Aminobenzoate degradation	3	1
Th1 and Th2 cell differentiation	3	1
T cell receptor signaling pathway	3	1
Phenylalanine metabolism	2	2
Tyrosine metabolism	2	2
One carbon pool by folate	1	1
Glyoxylate and dicarboxylate metabolism	1	1
Aminoacyl-tRNA biosynthesis	1	1
Tryptophan metabolism	1	1
Pyruvate metabolism	1	1
Ubiquinone and other terpenoid-quinone biosynthesis	1	1
Phenylalanine, tyrosine and tryptophan biosynthesis	1	1
Cyanoamino acid metabolism	1	1
Caffeine metabolism	1	1
Phosphatidylinositol signaling system	1	1
Biosynthesis of antibiotics	1	1
Glycosylphosphatidylinositol (GPI)-anchor biosynthesis	1	1
Cysteine and methionine metabolism	1	1
Folate biosynthesis	1	1
Inositol phosphate metabolism	1	1
Glycine, serine and threonine metabolism	1	1
Drug metabolism - other enzymes	1	1
Alanine, aspartate and glutamate metabolism	1	1
Valine, leucine and isoleucine degradation	1	1
Lipoic acid metabolism	1	1
Methane metabolism	1	1
Other glycan degradation	1	1
Isoquinoline alkaloid biosynthesis	1	1

Table S2.4 Canonical pathways identified via KEGG in the *peak infection* phase

Canonical pathway name	No. Sequences	No. Enzymes
Purine metabolism	5	4
Aminobenzoate degradation	4	1
Thiamine metabolism	4	1
Th1 and Th2 cell differentiation	4	1
T cell receptor signaling pathway	4	1
Metabolism of xenobiotics by cytochrome P450	2	2
One carbon pool by folate	1	1
Glyoxylate and dicarboxylate metabolism	1	1
Mannose type O-glycan biosynthesis	1	1
Pentose and glucuronate interconversions	1	1
Tryptophan metabolism	1	1
Histidine metabolism	1	1
Glycosaminoglycan biosynthesis - heparan sulfate / heparin	1	1
Cyanoamino acid metabolism	1	1
Caffeine metabolism	1	1
Phosphatidylinositol signaling system	1	1
Biosynthesis of antibiotics	1	1
Glycosylphosphatidylinositol (GPI)-anchor biosynthesis	1	1
Steroid hormone biosynthesis	1	1
Folate biosynthesis	1	1
Glycosaminoglycan biosynthesis - chondroitin sulfate / dermatan sulfate	1	1
Inositol phosphate metabolism	1	1
Glycine, serine and threonine metabolism	1	1
Ascorbate and aldarate metabolism	1	1
Drug metabolism - cytochrome P450	1	1
Retinol metabolism	1	1
Methane metabolism	1	1
Porphyrin and chlorophyll metabolism	1	1
Fructose and mannose metabolism	1	1

Table S2.5 Canonical pathways identified via KEGG in the *recovery* phase

Canonical pathway name	No. Sequences	No. Enzymes
Purine metabolism	15	10
Thiamine metabolism	6	1
Aminobenzoate degradation	4	1
Pyrimidine metabolism	4	4
Steroid hormone biosynthesis	3	2
Phosphatidylinositol signaling system	3	3
Drug metabolism - other enzymes	3	2
Th1 and Th2 cell differentiation	3	1
T cell receptor signaling pathway	3	1
Drug metabolism – cytochrome P450	2	1
Retinol metabolism	2	1
Metabolism of xenobiotics by cytochrome P450	2	1
Pentose and glucuronate interconversions	2	1
Ascorbate and aldarate metabolism	2	1
Porphyrin and chlorophyll metabolism	2	1
Mannose type O-glycan biosynthesis	1	1
Glutathione metabolism	1	1
Glycosaminoglycan biosynthesis - heparan sulfate / heparin	1	1
Betalain biosynthesis	1	1
Caffeine metabolism	1	1
Biosynthesis of antibiotics	1	1
Glycosylphosphatidylinositol (GPI)-anchor biosynthesis	1	1
Nicotinate and nicotinamide metabolism	1	1
Glycosaminoglycan biosynthesis - chondroitin sulfate / dermatan sulfate	1	1
Lipoic acid metabolism	1	1
One carbon pool by folate	1	1
Phenylalanine metabolism	1	1
Sphingolipid metabolism	1	1
Aminoacyl-tRNA biosynthesis	1	1
Tropane, piperidine and pyridine alkaloid biosynthesis	1	1
Histidine metabolism	1	1
Glycerophospholipid metabolism	1	1
Inositol phosphate metabolism	1	1
Carbon fixation pathways in prokaryotes	1	1
beta-Alanine metabolism	1	1
Glycine, serine and threonine metabolism	1	1
Fructose and mannose metabolism	1	1
Isoquinoline alkaloid biosynthesis	1	1

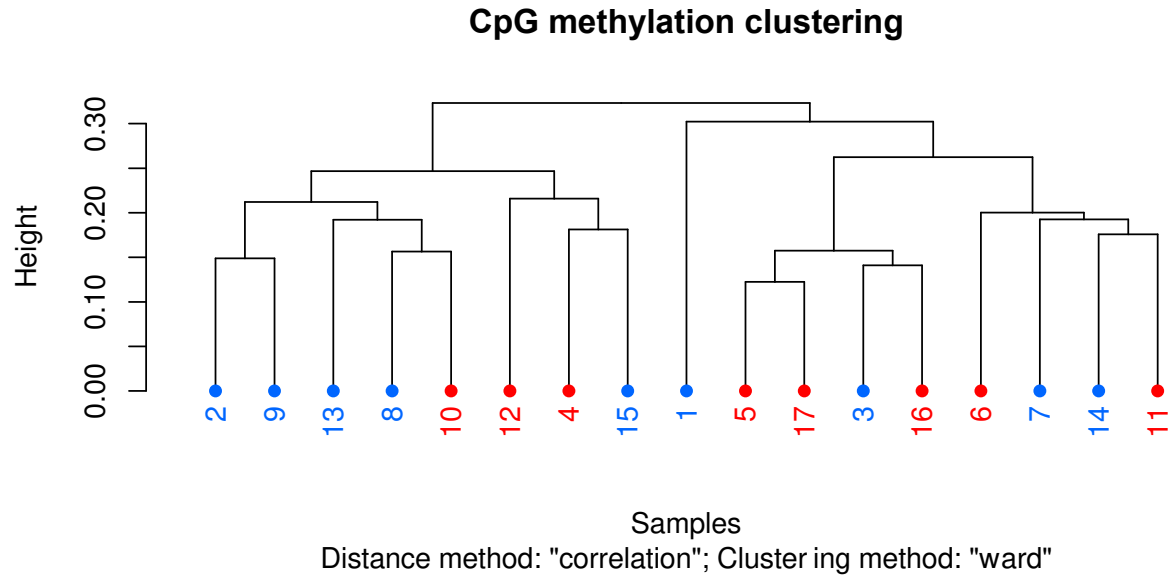


Fig. S2.1 Dendrogram of genome-wide CpG methylation value for all 17 samples. Samples are numbered and colour-coded with their experimental groups (*infected* fish are blue, *control* fish are red). Height is the Euclidean distance after hierarchical clustering of samples. 1-3: *infected* fish in *early infection* phase; 4-6: *control* fish in *early infection* phase; 7-9: *infected* fish in *peak infection* phase; 10-12: *control* fish in *peak infection* phase; 13-15: *infected* fish in *recovery* phase; 16-17: *control* fish in *recovery* phase.

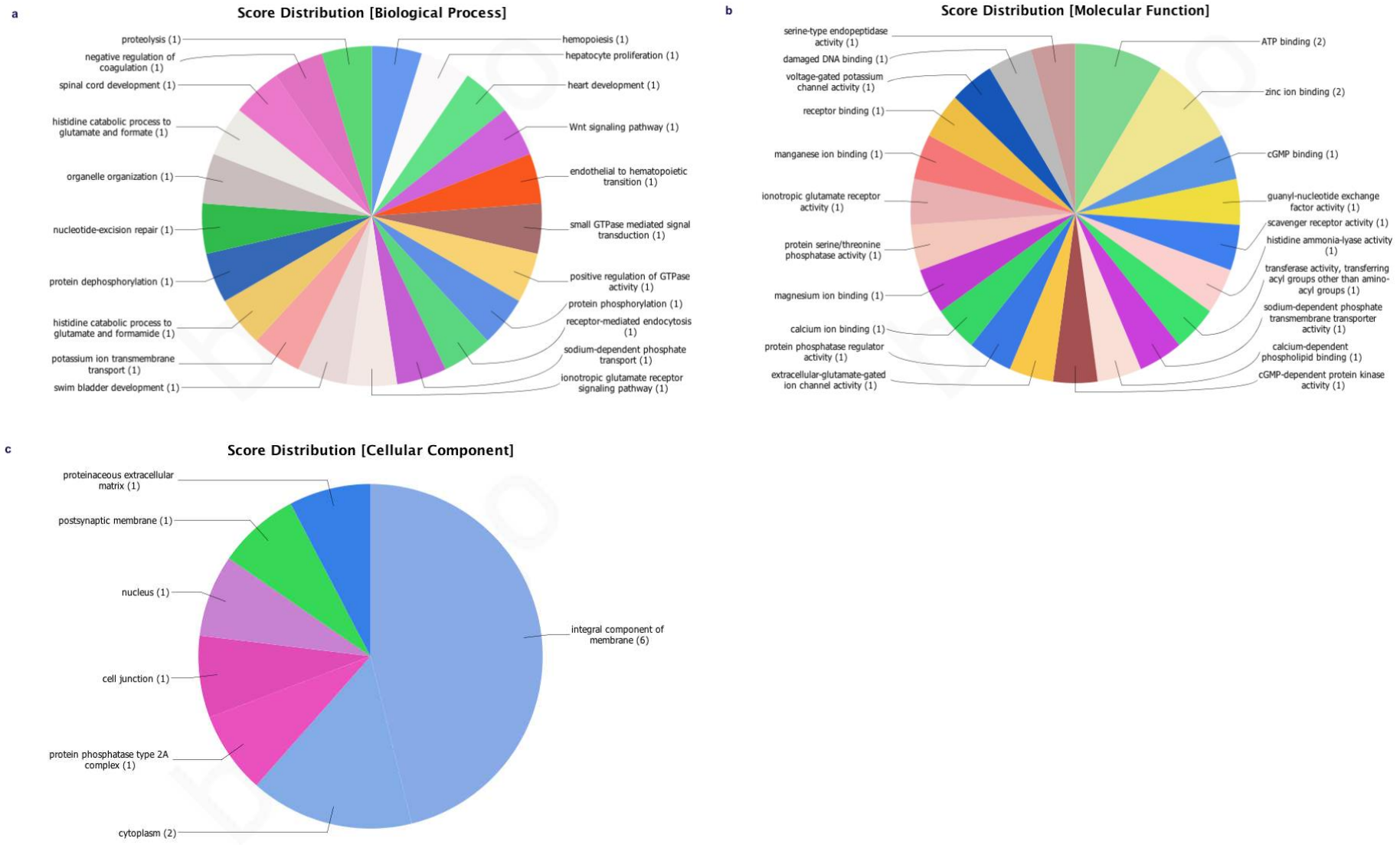


Fig. S2.2 Gene ontology node score distribution for the DMRs in *infected* vs. *control* guppies across all phases. The number in the parenthesis is the node score of its corresponding GO term. (a) Biological processes, (b) Molecular functions, (c) Cellular components.

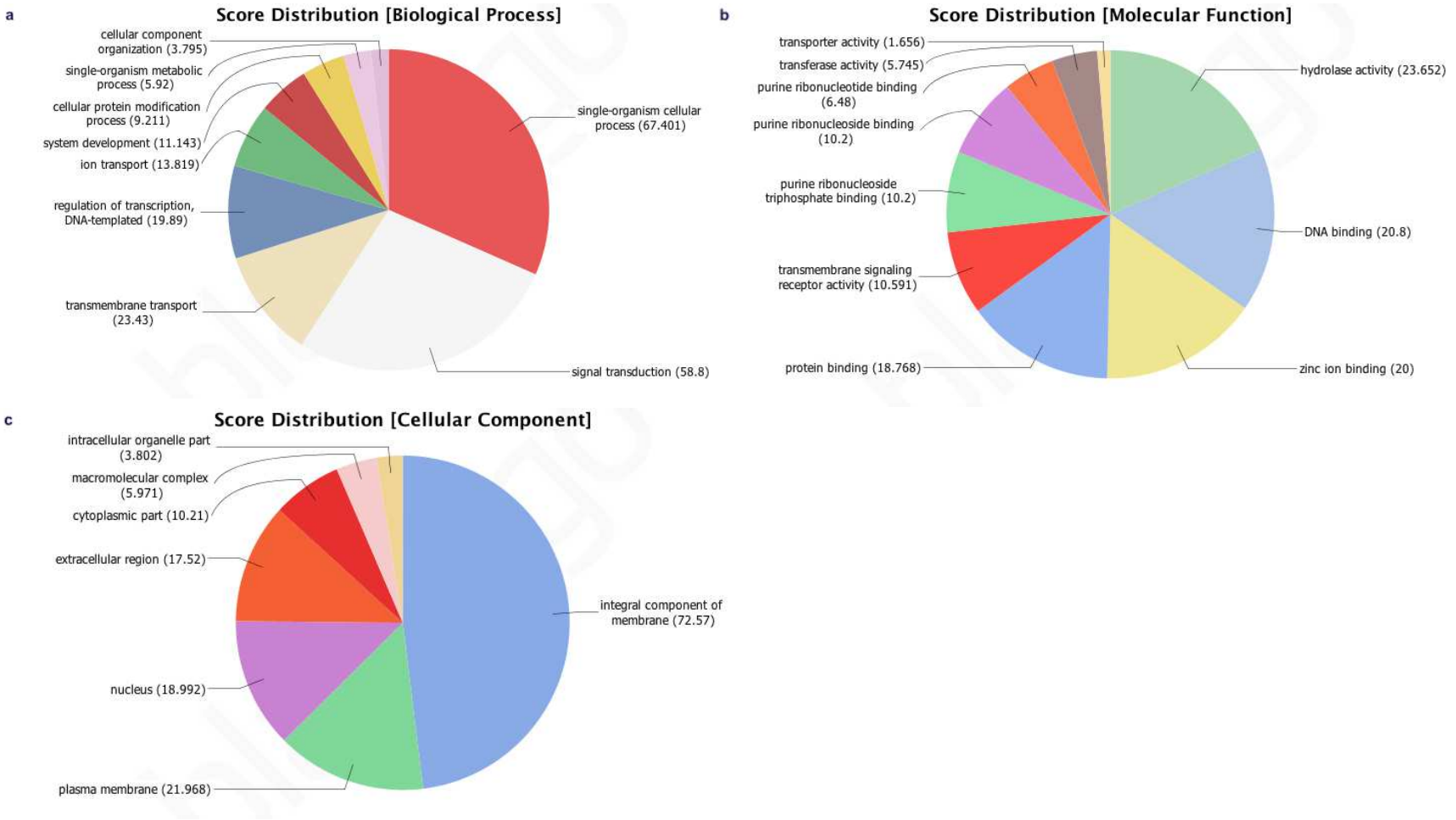


Fig. S2.3 Gene ontology node score distribution for the DMRs in the *early infection* phase. The number in the parenthesis is the node score of its corresponding GO term. (a) Biological processes, (b) Molecular functions, (c) Cellular components.

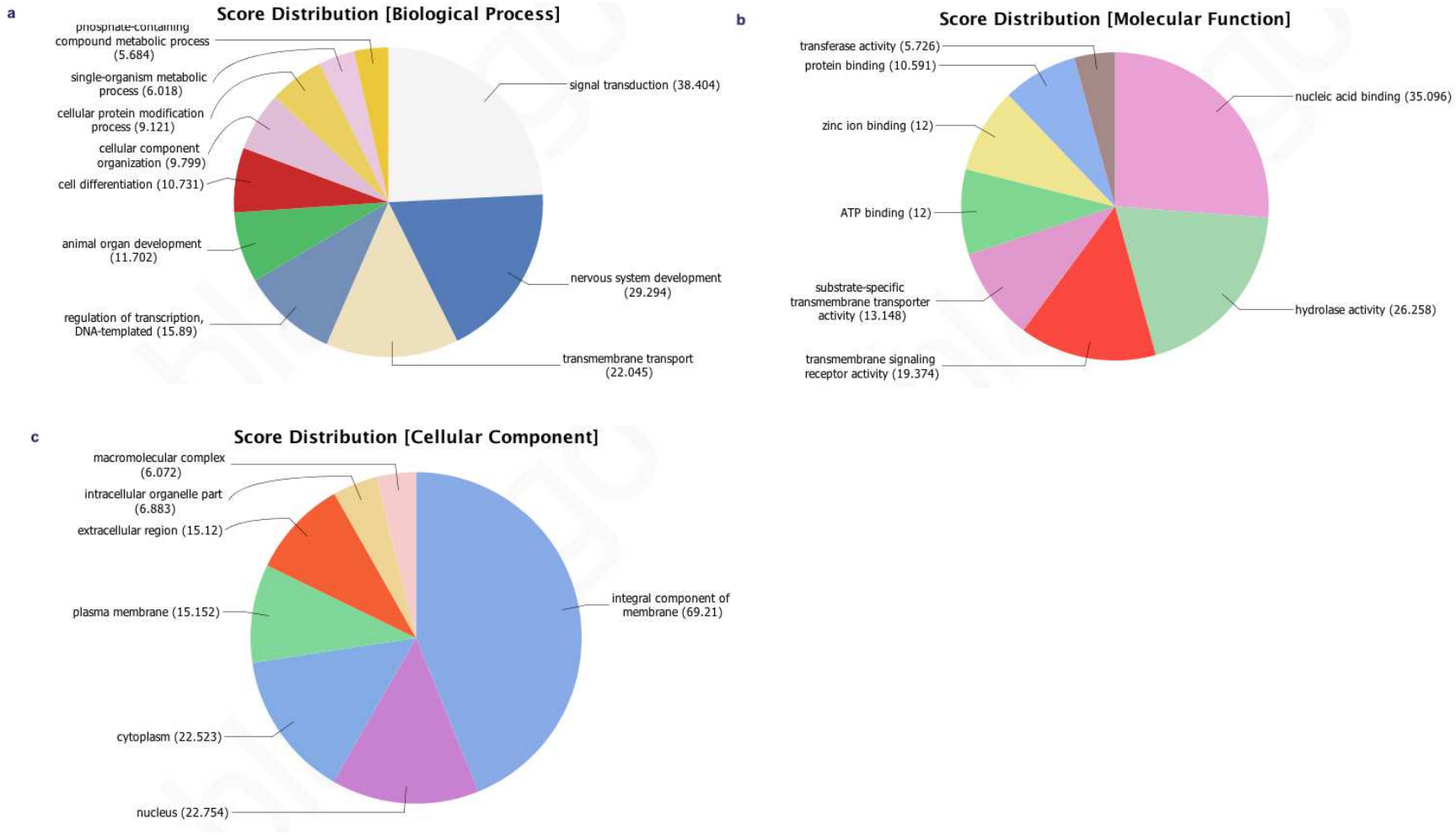


Fig. S2.4 Gene ontology node score distribution for the DMRs in the *peak infection* phase. The number in the parenthesis is the node score of its corresponding GO term. (a) Biological processes, (b) Molecular functions, (c) Cellular components.

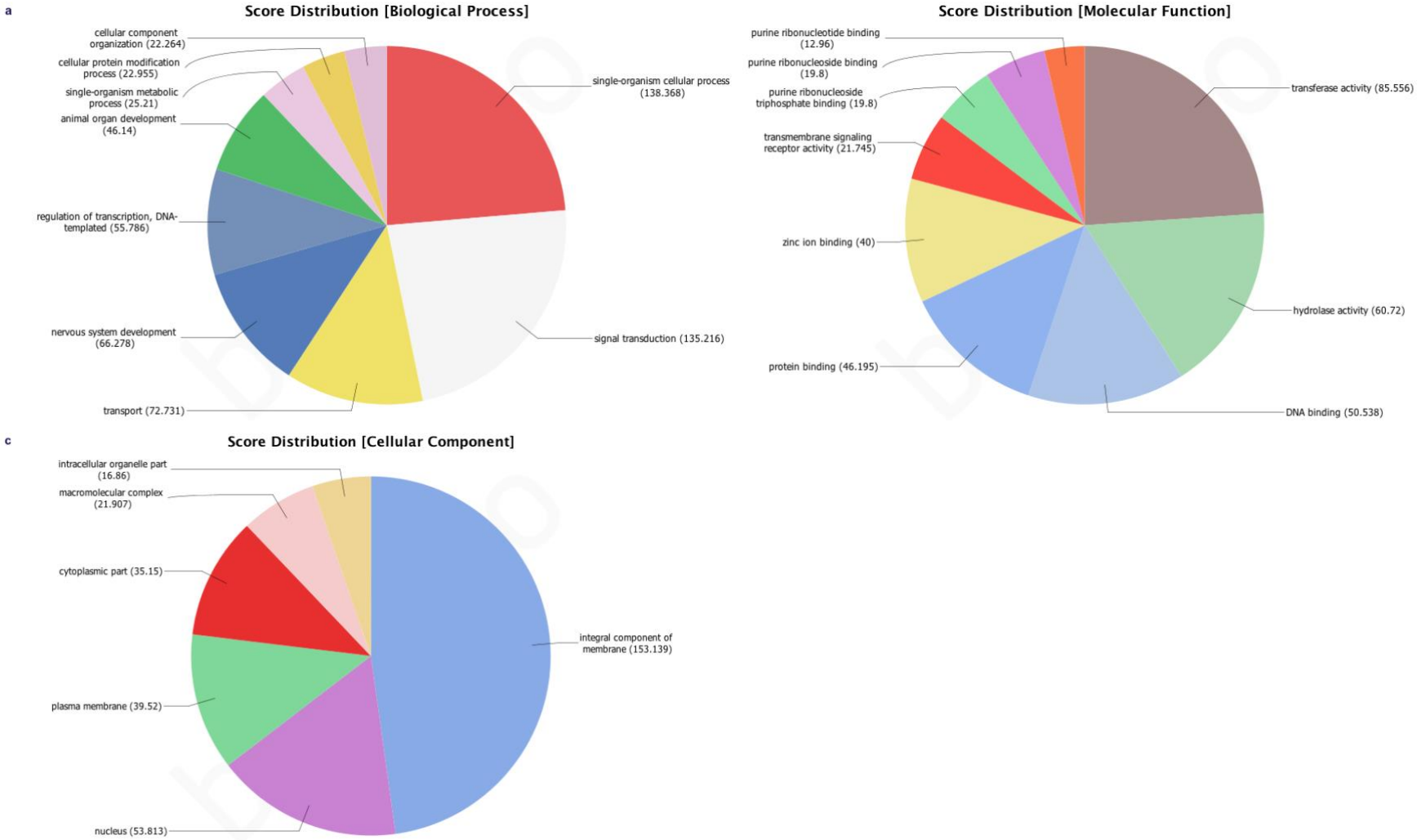


Fig. S2.5 Gene ontology node score distribution for the DMRs in the *recovery* phase. The number in the parenthesis is the node score of its corresponding GO term. (a) Biological processes, (b) Molecular functions, (c) Cellular components.

CONNECTING STATEMENT

In Chapter 2, I analysed parasite-induced epigenetic variation in guppies across distinct infection phases, and found hypermethylation as a general response to parasite infection. I also found a unique DNA methylation pattern in each infection phase, with several of the differentially methylated regions overlapping with immune genes. These results indicate that, in addition to endoparasites, ectoparasites can also trigger genome-wide epigenetic responses, and that epigenetic modifications can be an important mechanism facilitating host resistance to parasite infection. However, environmental complexity is usually higher in nature than in the lab, and thus epigenetic responses to the full suite of environmental stressors experienced in nature may be different those found in lab-reared animals.

To test the effects of natural levels of environmental complexity on epigenetic variation, in Chapter 3, I characterise DNA methylation patterns at the early stage of colonization, using lizard populations transplanted to islands that vary in their habitat quality based on several environmental parameters. Previous studies have typically focused on phenotypic and genetic variation shaped by long-term adaptation to distinct environment types, whereas the mechanisms underlying shorter-term responses to environment change remain unclear. Thus, results from Chapter 3 will improve our understanding of how epigenetic modifications facilitate responses to novel environments at the earliest stages of colonization.

CHAPTER 3: An epigenetic signature of colonizing new environments

Authors: Juntao Hu[*,1], Arash M. Askary[1], Timothy J. Thurman[1,2], Dave Spiller[3], Todd M. Palmer[4,5], Robert M. Pringle[6], and Rowan D.H. Barrett[1]

[1]Redpath Museum and Department of Biology, McGill University, 859 Sherbrooke St. W., Montreal, Quebec, Canada H3A 0C4
[2]Smithsonian Tropical Research Institute. Panamá, República de Panamá.
[3]Section of Evolution and Ecology and Center for Population Biology, One Shields Avenue, University of California, Davis, California USA 95616-8755
[4]Mpala Research Centre, P.O. Box 555, Nanyuki Kenya 10400
[5]Department of Biology, University of Florida, Gainesville, Florida USA 32611
[6]Department of Ecology & Evolutionary Biology, Princeton University, Princeton, New Jersey USA 08544

Keywords: DNA methylation, epigenomics, phenotypic plasticity, *Anolis sagrei*, reduced representation bisulphite sequencing

[*]Corresponding author: Juntao Hu, juntao.hu@mail.mcgill.ca

Abstract

Founder populations often show rapid divergence from source populations after colonizing new environments. Epigenetic modifications can mediate phenotypic response to environmental change and may be an important mechanism promoting such rapid differentiation in founder populations. Whereas many long-term studies have explored the extent to which divergence between source and founder populations is genetically heritable versus plastic, the role of epigenetic processes during colonization events remains unclear. To investigate epigenetic modifications in founding populations, we experimentally colonized eight natural Caribbean islands with brown anole lizards (*Anolis sagrei*) from a common source population; we then quantitatively measured genome-wide DNA methylation in liver tissue, using reduced representation bisulphite sequencing (RRBS), of individuals transplanted onto islands with high- versus low-habitat quality. We detected significant effects of transplantation, with an average of 24.8% differential methylation per differentially methylated region (DMR). We also found distinct epigenetic responses to colonization of high- versus low-quality islands, with most DMRs mapping to genes thought to encode proteins with functions likely to be relevant to habitat change (e.g., neuronal modification, skeletal and myocardial growth, immune response). This study provides the first experimental evidence of a relationship between epigenetic variation and the earliest stages of colonization of novel environments, and suggests that habitat quality can play an important role in governing the nature of these epigenetic modifications.

Introduction

Island populations often rapidly diverge from their relatives on nearby islands or mainlands, a phenomenon with profound ecological and evolutionary implications. Such genetic and/or phenotypic differentiation can result from genetic drift occurring in small founding populations, or from island environments differing from each other and from the mainland, which can impose divergent natural selection (Kolbe et al. 2012; Wessel et al. 2013). Previous work has typically focused on the long-term consequences of founder effects and natural selection in colonizing populations over multiple generations, whereas the mechanisms underlying short-term phenotypic response to colonization of new environments remain relatively unexplored (Losos et al. 1994; Schoener et al. 2001; Kolbe et al. 2012). Recently, both empirical and theoretical studies have demonstrated that epigenetic modifications, particularly DNA methylation, can contribute to rapid phenotypic changes by modulating the gene regulatory response to environmental conditions (Feil and Fraga 2012). For example, studies of DNA methylation have revealed associations with several environmentally responsive phenotypes, such as distinct castes in eusocial insects caused by divergent nutritional intake (Kucharski et al. 2008), sex-ratio shifts in turtles related to distinct early developmental temperatures (Matsumoto et al. 2013), and alteration of muscle composition in fish caused by different rearing temperatures (Campos et al. 2013). Recent work has suggested that expansion into new environments can lead to widespread changes in DNA methylation levels (McErlean et al. 2014; Baerwald et al. 2015; Lea et al. 2016). However, these studies have compared long-established populations (up to 25 generations) with populations from what is thought to be the original source habitat, and thus cannot determine how early in the colonization process these changes may have occurred. Here, we report the results of a replicated whole-ecosystem field experiment in which we transplanted individuals onto small islands of differing habitat quality in order to quantify the epigenetic modifications accompanying the very earliest (within generation) stage of colonization of new environments by founder populations.

Due to their experimental tractability and a rich history of studies documenting population dynamics during colonization of small islands, Caribbean *Anolis* lizards are an ideal model system for the study of many ecological and evolutionary phenomena (Losos 2009). For example, there is evidence that populations can undergo relatively rapid differentiation in fitness-related traits (e.g., length of hindlimbs) due to both founder effects and natural selection (Losos et al. 1997; Losos 2009; Kolbe et al. 2012; Jaffe et al. 2016). In

addition, population size has been shown to be positively correlated with vegetated island area, both in populations transplanted from source islands to small experimental islands (Kolbe et al. 2012) and in populations naturally re-colonizing small islands where previous populations were extirpated by hurricanes (Schoener et al. 2001). However, the molecular mechanisms that might underpin immediate responses of lizards to new environments remain unexplored, and no study to date has investigated epigenetic modifications of any kind in *Anolis* lizards. Here, we provide the first single-nucleotide-resolution measurement of DNA methylation in *A. sagrei* and use experimental introductions of lizards to multiple small islands in the vicinity of a source island to evaluate the influence of habitat quality on epigenetic responses at the earliest stage of island colonization. We predicted that the magnitude of epigenetic response in our experiment (measured as differential methylation between the source population and colonizing populations) would increase with the extent of the environmental difference between the source island and newly colonized islands. Furthermore, because it has been suggested that responses to new environments over short time periods may be predictive of large-scale patterns on longer timescales (Losos et al. 1997; Arnold et al. 2001; Simons 2002), we predict that differentially methylated regions will map to genes with functions likely to be important for survival and fitness (e.g., skeletal growth; Kolbe et al. 2012) in novel environments.

Results

To assess the epigenetic responses of lizards to new environments, we introduced 219 lizards (84 males and 135 females) from the source island, Staniel Cay (Exumas, Bahamas; Fig. 3.1) to multiple small islands within 10 km. Prior to introductions, we surveyed each island to collect vegetation and climatic data, which we used to cluster islands into high- and low-quality groups (Fig. 3.2; Supplementary text S3.1, Fig. S3.1), as well as to verify the absence of any existing lizard population. Low-quality islands had an average of 28.4% less vegetated area (hereafter, "area"; means ± SD: low-quality = 235 ± 200 m^2, high-quality = 328 ± 129 m^2) and 75.6% greater diel temperature fluctuation than did high-quality islands (low-quality = 23.7 ± 2.9 ºC, high-quality = 13.5 ± 1.1 ºC). All experimental islands were much smaller and more thermally variable than the source island: the high-quality experimental islands had an average of 99.9% less area (source = 1,385,637 m^2) and 57.0% greater diel temperature fluctuation (source = 8.6 ± 1.2 ºC) than the source. We expected the distinct patterns of

environmental variation between our treatments to cause divergent levels of physiological stress to lizards, and thus to induce different patterns of genome-wide methylation.

We recaptured lizards from each experimental island after exposing transplanted lizards to the new environment for at least four days. To explore short-term epigenetic responses to this earliest stage of colonization, we performed three comparisons: high- versus low-quality islands, high-quality islands versus the source island, and low-quality islands versus the source island. We identified 275 (high- vs. low-quality), 667 (high-quality vs. source), and 1116 (low-quality vs. source) differentially methylated cytosines (DMCs) (Fig. 3.3). Permutation analyses (Supplementary text S3.3) revealed that among these comparisons, only the comparison between low-quality islands and the source island had a significantly greater number of DMCs than expected by chance ($P = 0.026$, compared with $P = 0.144$ for high- vs. low-quality and $P = 0.195$ for high-quality vs. source; Fig. S3.2). Among these significant DMCs between low-quality islands and the source island, we found no significant difference in the number of cytosines that were hypermethylated versus hypomethylated in lizards from the low-quality islands relative to the source island (574 vs. 542, $\chi^2 = 0.91$, df = 1, $P = 0.34$; Fig. 3.3), suggesting a broad range of environmental causes underlying this general methylation response. These results are consistent with our hypothesis that colonization of new environments can lead to genome-wide changes in DNA methylation, with the magnitude of this response being dependent on the extent of the environmental shift between source and colonized habitats. Notably, we did not detect significant effects of sex on DNA methylation levels throughout our analyses, which is consistent with previous studies that have identified weak or no sex effects on methylation differences between distinct environments (Lea et al. 2016). We also found no significant differences between the four individual islands within each of our habitat-quality treatments.

On the basis of DMCs, we identified nine differentially methylated regions (DMRs) between lizards from low-quality islands and the source island, after correcting for false discovery rate (Table 3.1). To be considered a DMR, a region needed to contain at least three CpG sites within a genomic distance defined by a bimodal normal distribution fitting on the log2 distance of adjacent CpGs across the genome (Li et al. 2013), with at least one classified as a DMC and an absolute mean methylation difference > 20% when comparing lizards from different island categories. Above this 20% lower-limit threshold, methylation differences between the source and low-quality islands ranged from 21% to 36% per DMR, with a mean methylation difference (± SD) of 25.38% ± 5.32% per DMR. Due to the incomplete

annotation of the *A. sagrei* genome, we were not able to extract the genomic context (i.e., promoters, exons, introns, intergenic regions) for each DMR, making it difficult to predict the effects of differential methylation on gene expression. With a few exceptions, individuals clustered primarily by their island category (low-quality or source) based on the DMRs (Fig. 3.4). By blasting against the NCBI non-redundant database, we found that five of the nine DMRs were annotated with known genes that have functions relevant for responding to environmental change (Table 3.1, see discussion below), while the remaining DMRs were without significant hits on known genes.

Discussion

Phenotypically plastic responses play a central role in facilitating the establishment success of animal populations colonizing new habitats (Lande 2015), which in turn influences key ecological and evolutionary processes such as biological invasion and adaptive radiation. We experimentally simulated a natural colonization scenario on previously unoccupied oceanic islands and used a high-resolution technique (RRBS) to quantitatively measure DNA methylation across the *A. sagrei* genome, revealing that colonization of low-quality environments can lead to significant changes in methylation levels. Some of the DMRs mapped to genes that are plausibly relevant to survival and fitness in novel environments (see below), suggesting the possibility that epigenetic mechanisms influence the expression of these genes and associated phenotypic plasticity. Our results are consistent with other recent studies demonstrating that habitat quality can affect methylation levels (Huang et al. 2017; Le Luyer et al. 2017; McNew et al. 2017). To our knowledge, this study represents the first evidence of methylation responses in populations at the earliest stage of colonizing new environments (i.e., within 4–13 days of colonization).

We found that the number of DMRs between populations on colonized islands and the source island was 1.7 times greater when lizards were moved to islands with low-quality versus high-quality habitat. Indeed, the number of DMRs detected in the high-quality island versus source comparison was low enough that it could not be distinguished from chance, suggesting that not all environmental change will necessarily result in sufficient stress to trigger an epigenetic response. Our high-quality islands are likely to have represented a modest enough shift in conditions that lizards did not require plastic responses for acclimation (at least in the short term). In contrast, the highly variable temperature and lack of vegetated area encountered by lizards on low-quality islands presented a significant

environmental challenge that led to widespread changes in methylation relative to source populations. In addition to immediate effects on gene expression or repression, changes in methylation levels could also lead to accelerated mutation rates over evolutionary time scales, due to the positive relationship between the rate of deamination of methylated cytosine and genomic DNA methylation levels (Roberts and Gavery 2012). Thus, it is possible that altered methylation levels in response to short-term environmental change could ultimately result in greater genetic variation, and augment the phenotypic differentiation observed in longer-term studies of *A. sagrei* introduced onto new islands (Losos et al. 1997; Kolbe et al. 2012).

We mapped several DMRs to genes that have relevance to environmental change. For example, *TIP39* is involved in thermoregulation (Dimitrov et al. 2011) and pain sensing (Usdin et al. 2003). *LINO3* has influences on neuronal survival by regulating axon regeneration and oligodendrocyte maturation (Stefansson et al. 2009). *IGFBP5* is involved in osteoblast differentiation and skeletal growth (Mukherjee and Rotwein 2008). *ALPK3* plays a role in establishing normal structure and functions of the myocardium (Van Sligtenhorst et al. 2011). *SIGLEC1* can contribute to inflammatory processes, and is related to disease resistance (Eakin et al. 2016). This collection of genes may be important for facilitating the colonization of environments with novel sensory, physical, or immune challenges. Recent studies have shown that the selective agents at the early stages of colonization (e.g., novel food resources, temperature regime, and parasites) can trigger physiological responses affecting skeletal growth (Ramirez et al. 2015), cardiorespiratory physiology and myocardium functions (Borrelli et al. 2008; Eliason et al. 2011; McBryan et al. 2013), and immune response (Acevedo-Whitehouse and Duffus 2009). Among the five genes above, three were hypomethylated (*LINGO3*, *ALPK3*, *SIGLEC1*), and two were hypermethylated (*IGFBP5*, *TIP39*), suggesting that distinct mechanisms regulating the activation and repression of genes might work in concert to facilitate lizards responding to the new and stressful environments. Although we can only speculate about the potential phenotypic implications of hyper- and hypomethylation of these genes at present, our work contributes to a foundation for future work aimed at testing such linkages.

The genome-wide methylation data presented here provide the first investigation into the relationship between epigenetic variation and response to short-term environmental change at the very earliest stage of (experimentally controlled) colonization in natural animal populations. It also demonstrates the power of a high-resolution sequencing technique for identifying loci potentially associated with complex cellular responses under natural settings. Our results indicate that consistent changes in DNA methylation can occur within only a few

days of immersion in a novel environment, and may thus be an important molecular mechanism for regulating responses to environmental stressors during the colonization of novel habitats.

Table 3.1 Differentially methylated regions (DMRs) identified in low-quality vs. source

Symbol/Scaffold ID	Entrez gene name	Meth diff‡	*P*-value	*Q*-value
Low-quality vs. source				
IGFBP5	Insulin-like growth factor binding-protein 5	21.4	2.3E-15	2.6E-15
LINGO3	Leucine-rich repeat and immunoglobulin-like domain-containing nogo receptor-interacting protein 3	-20.6	1.3E-29	3.1E-29
ScPFAeV_13729:HRSCAF=14524	Unknown	23.2	8.2E-22	1.3E-21
ALPK3	Alpha kinase 3	-21.2	2.9E-21	4.3E-21
ScPFAeV_13733:HRSCAF=14596	RNA-directed DNA polymerase	-32.2	7.6E-36	4.8E-35
ScPFAeV_7291:HRSCAF=7759	Unknown	-35.9	4.8E-47	5.2E-46
ScPFAeV_8734:HRSCAF=9256	Unknown	22.7	6.6E-22	1.1E-21
TIP39	Tuberoinfundibular peptide of 39 residues	24.8	4.6E-19	6.5E-19
SIGLEC1	Sialoadhesin	-26.4	8.9E-57	3.9E-55

‡Percent methylation differences averaged from all CpG sites within the defined region. Positive values represent increased methylation in low-quality island samples, and negative values represent increased methylation in source island samples.

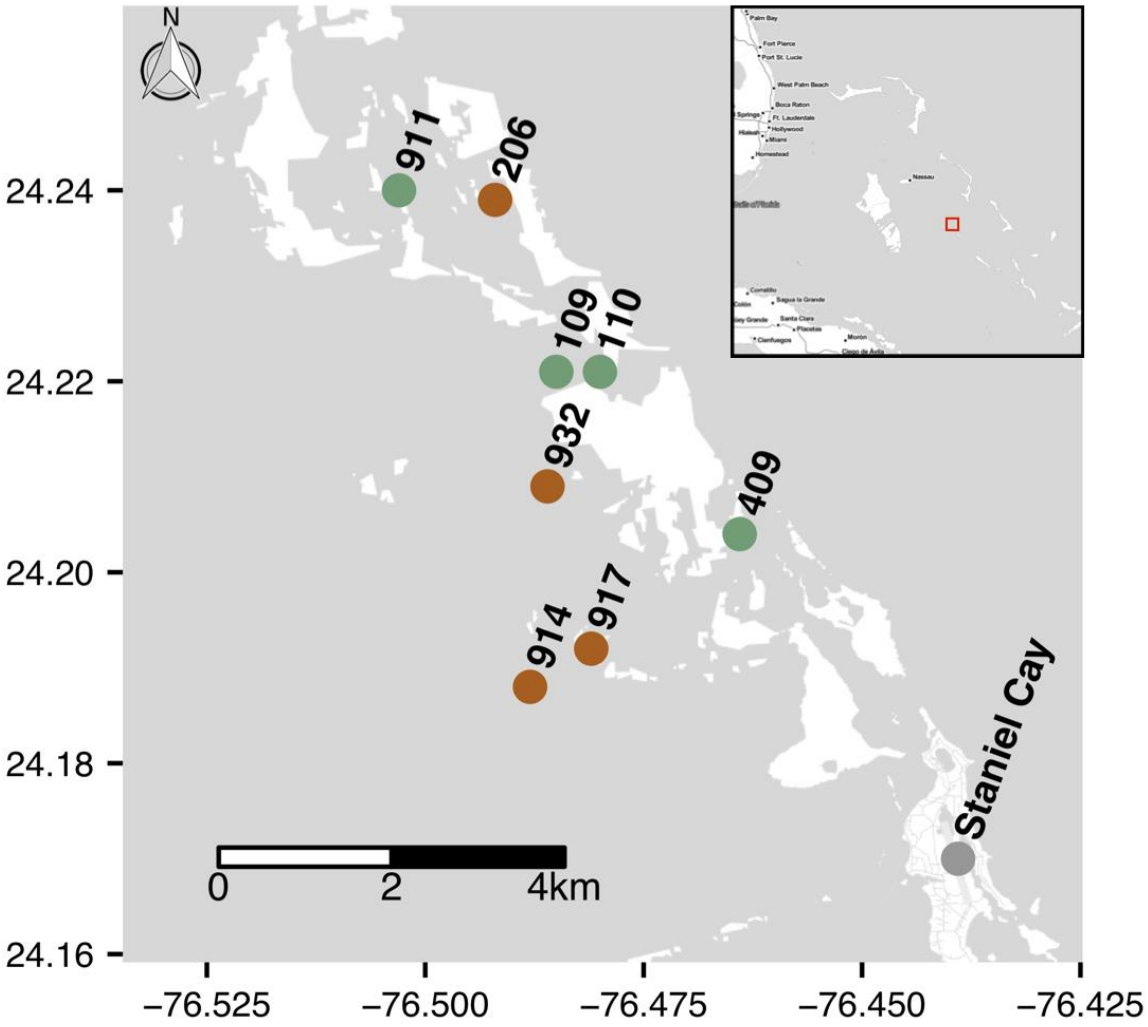


Fig. 3.1 Map of treatment islands. High-quality islands are shown in green, low-quality islands are shown in orange. Source island is shown in grey. Red square in the inset shows location of experiment in relation to the broader geographic region (the Exuma Chain of the Bahamas).

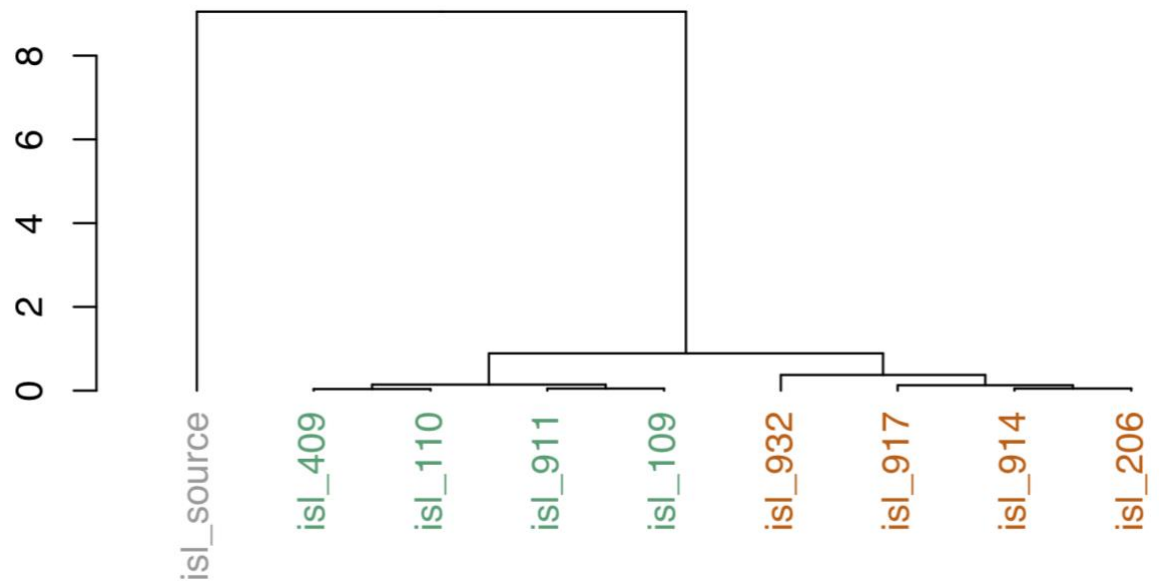


Fig. 3.2 Dendrogram of island quality for all treatment islands. Height is the Euclidean distance after hierarchical clustering of islands, based on scaled vegetated area, and the scaled temperature difference between maximum and minimum temperatures of each island. High-quality islands are shown in green, low-quality islands are shown in orange, the source island is shown in grey.

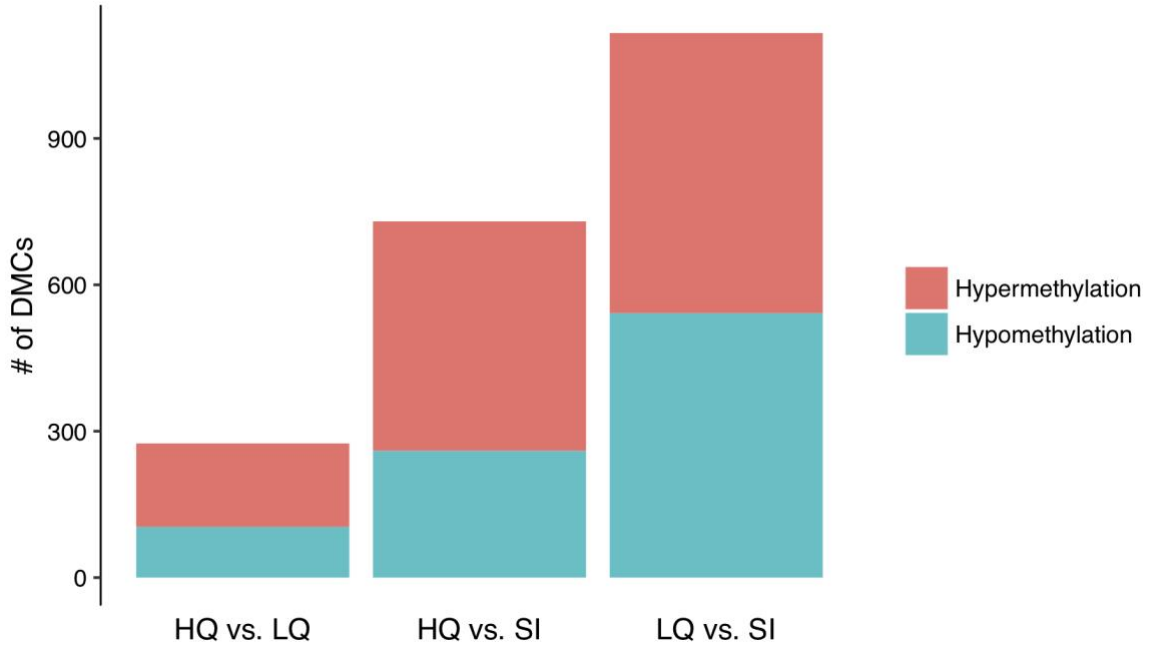


Fig. 3.3 Number of CpG sites that are hypomethylated (blue) and hypermethylated (red) in the three comparisons (high-quality islands, HQ; low-quality islands, LQ; the source island, SI). The y-axis reflects the number of hypo- and hypermethylated DMCs for the former island category relative to the latter island category. For example, in HQ vs. LQ, the blue and red bars reflect the number of hypo- and hypermethylated DMCs for HQ relative to LQ.

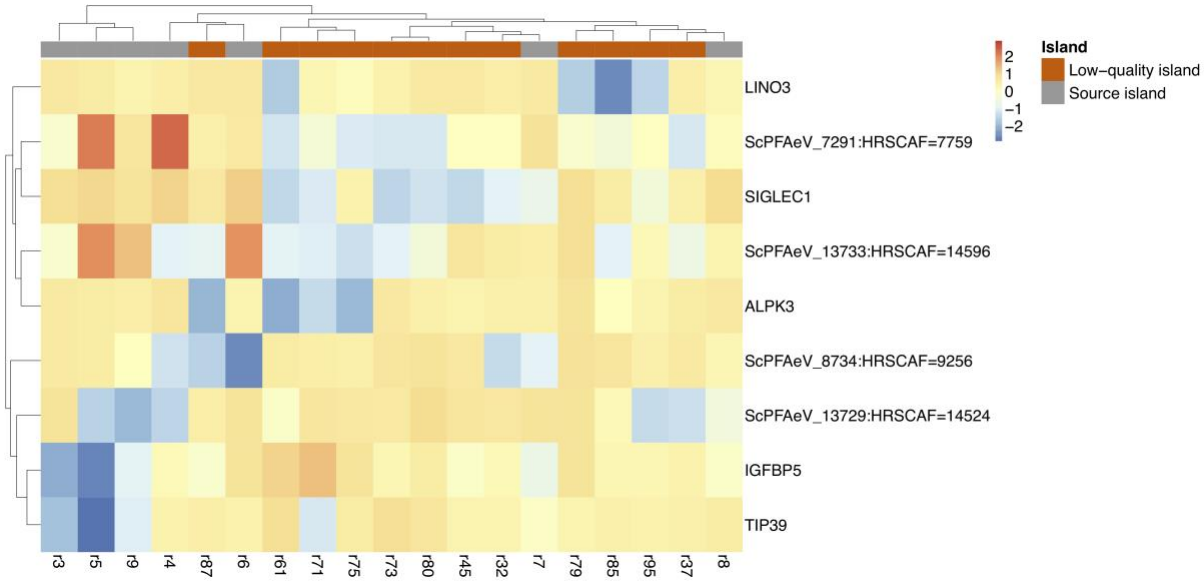


Fig. 3.4 Heatmap of methylation levels of DMRs in low-quality vs. the source island. Each column represents a single individual, colored according to its island-type affiliation. Each row represents one of the DMRs, which are clustered based on the similarities of the methylation patterns among individuals. Darker orange indicates greater methylation in an individual for that DMR. Darker blue indicates lesser methylation in an individual for that DMR. Individual dendrogram positions are based on overall methylation patterns across DMRs. The gene names or the Contigs IDs that DMRs were mapped to are shown on the corresponding rows.

Acknowledgements The authors would like to thank John and Dreko Chamberlain for their assistance with various facets of the study, and the Bahamas Government for its permission to conduct this research. Stacy Lubin-Gray of the Bahamas Environment, Science, and Technology (BEST) Commission, Luceta Hanna of the Bahamas Department of Agriculture, and their colleagues provided invaluable assistance with permitting.

Ethics This work was conducted in accordance with Bahamian law, permits from BEST Commission and Bahamas Department of Agriculture, and Princeton University's Institutional Animal Care and Use Committee.

Data accessibility All data can be found as electronic supplementary material accompanying this manuscript.

Authors' contributions A.A., J.H., and R.D.H.B. conceived and designed the study; A.A., J.H., T.J.T., D.A.S., T.M.P., R.M.P., and R.D.H.B. performed experiments; J.H. analyzed the data; R.M.P. and R.D.H.B. obtained funding; J.H. and R.D.H.B. wrote the manuscript with input from T.J.T. All authors agree to be held accountable for the work, and approve the final version to be published.

Competing interests We declare we have no competing interests.

Funding This work is a product of U.S. National Science Foundation grant DEB-1457691 to R.M.P. and R.D.H.B, a Natural Science and Engineering Research Council of Canada Discovery Grant and Canada Research Chair to R.D.H.B., and State Scholarship Fund Award, China Scholarship Council 201406350023 to J.H.

References

Acevedo-Whitehouse K, Duffus ALJ. 2009. Effects of environmental change on wildlife health. *Philos T R Soc B* 364:3429-3438.

Arnold SJ, Pfrender ME, Jones AG. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Genetica* 112:9-32.

Borrelli E, Nestler EJ, Allis CD, Sassone-Corsi P. 2008. Decoding the epigenetic language of neuronal plasticity. *Neuron* 60:961-974.

Campos C, Valente LMP, Conceicao LEC, Engrola S, Fernandes JMO. 2013. Temperature affects methylation of the myogenin putative promoter, its expression and muscle cellularity in Senegalese sole larvae. *Epigenetics* 8:389-397.

Dimitrov EL, Kim YY, Usdin TB. 2011. Regulation of hypothalamic signaling by tuberoinfundibular peptide of 39 residues is critical for the response to cold: a novel peptidergic mechanism of thermoregulation. *J Neurosci.* 31:18166-18179.

Eakin AJ, Bustard MJ, McGeough CM, Ahmed T, Bjourson AJ, Gibson DS. 2016. Siglec-1 and -2 as potential biomarkers in autoimmune disease. *PROTEOMICS – Clinical Applications* 10:635-644.

Eliason EJ, Clark TD, Hague MJ, Hanson LM, Gallagher ZS, Jeffries KM, Gale MK, Patterson DA, Hinch SG, Farrell AP. 2011. Differences in thermal tolerance among sockeye salmon populations. *Science* 332:109.

Feil R, Fraga MF. 2012. Epigenetics and the environment: emerging patterns and implications. *Nat Rev Genet.* 13:97-109.

Huang X, Li S, Ni P, Gao Y, Jiang B, Zhou Z, Zhan A. 2017. Rapid response to changing environments during biological invasions: DNA methylation perspectives. *Mol Ecol.* 26:6621-6633.

Jaffe AL, Campbell-Staton SC, Losos JB. 2016. Geographical variation in morphology and its environmental correlates in a widespread North American lizard, *Anolis carolinensis* (Squamata: Dactyloidae). *Biological Journal of the Linnean Society* 117:760-774.

Kolbe JJ, Leal M, Schoener TW, Spiller DA, Losos JB. 2012. Founder effects persist despite adaptive differentiation: a field experiment with lizards. *Science* 335:1086-1089.

Kucharski R, Maleszka J, Foret S, Maleszka R. 2008. Nutritional control of reproductive status in honeybees via DNA methylation. *Science* 319:1827-1830.

Lande R. 2015. Evolution of phenotypic plasticity in colonizing species. *Mol Ecol.* 24:2038-2045.

Le Luyer J, Laporte M, Beacham TD, Kaukinen KH, Withler RE, Leong JS, Rondeau EB, Koop BF, Bernatchez L. 2017. Parallel epigenetic modifications induced by hatchery rearing in a Pacific salmon. *Proc Natl Acad Sci U S A.*

Lea AJ, Altmann J, Alberts SC, Tung J. 2016. Resource base influences genome-wide DNA methylation levels in wild baboons (*Papio cynocephalus*). *Mol Ecol.* 25:1681-1696.

Li S, Garrett-Bakelman FE, Akalin A, Zumbo P, Levine R, To BL, Lewis ID, Brown AL, D'Andrea RJ, Melnick A, et al. (Li2013 co-authors). 2013. An optimized algorithm for detecting and annotating regional differential methylation. *BMC Bioinformatics* 14:1-9.

Losos J. 2009. Lizards in an evolutionary tree: ecology and adaptive radiation of anoles: University of California Press, Berkeley, CA.

Losos JB, Irschick DJ, Schoener TW. 1994. Adaptation and constraint in the evolution of specialization of Bahamian *Anolis* lizards. *Evolution* 48:1786-1798.

Losos JB, Warheit KI, Schoener TW. 1997. Adaptive differentiation following experimental island colonization in *Anolis* lizards. *Nature* 387:70.

Matsumoto Y, Buemio A, Chu R, Vafaee M, Crews D. 2013. Epigenetic control of gonadal aromatase (*cyp19a1*) in temperature-dependent sex determination of red-eared slider turtles. *PLoS ONE* 8:e63599.

McBryan TL, Anttila K, Healy TM, Schulte PM. 2013. Responses to temperature and hypoxia as interacting stressors in fish: Implications for adaptation to environmental change. *Integr Comp Biol.* 53:648-659.

McNew SM, Beck D, Sadler-Riggleman I, Knutie SA, Koop JAH, Clayton DH, Skinner MK. 2017. Epigenetic variation between urban and rural populations of Darwin's finches. *BMC Evol Biol.* 17:183.

Mukherjee A, Rotwein P. 2008. Insulin-like growth factor-binding protein-5 inhibits osteoblast differentiation and skeletal growth by blocking insulin-like growth factor actions. *Mol Endocrinol.* 22:1238-1250.

Ramirez MD, Avens L, Seminoff JA, Goshe LR, Heppell SS. 2015. Patterns of loggerhead turtle ontogenetic shifts revealed through isotopic analysis of annual skeletal growth increments. *Ecosphere* 6:1-17.

Roberts SB, Gavery MR. 2012. Is there a relationship between DNA methylation and phenotypic plasticity in invertebrates? *Front Physiol.* 2:116.

Schoener TW, Spiller DA, Losos JB. 2001. Natural restoration of the species-area relation for a lizard after a hurricane. *Science* 294:1525-1528.

Simons AM. 2002. The continuity of microevolution and macroevolution. *J Evol Biol.* 15:688-701.
Stefansson H, Steinberg S, Petursson H, Gustafsson O, Gudjonsdottir IH, Jonsdottir GA, Palsson ST, Jonsson T, Saemundsdottir J, Bjornsdottir G, et al. 2009. Variant in the sequence of the *LINGO1* gene confers risk of essential tremor. *Nat Genet.* 41:277-279.
Usdin TB, Dobolyi A, Ueda H, Palkovits M. 2003. Emerging functions for tuberoinfundibular peptide of 39 residues. *Trends in Endocrinology & Metabolism* 14:14-19.
Van Sligtenhorst I, Ding ZM, Shi ZZ, Read RW, Hansen G, Vogel P. 2011. Cardiomyopathy in α-kinase 3 (*ALPK3*)-deficient mice. *Vet Pathol.* 49:131-141.
Wessel A, Hoch H, Asche M, von Rintelen T, Stelbrink B, Heck V, Stone FD, Howarth FG. 2013. Founder effects initiated rapid species radiation in Hawaiian cave planthoppers. *Proceedings of the National Academy of Sciences* 110:9391

Supplementary text

S3.1. Experimental design

We collected 227 *A. sagrei* lizards from a source island, Staniel Cay, Bahamas (vegetated area 1,385,637 m^2), and transplanted individuals to treatment islands of contrasting environmental quality (Fig. 3.1; Table S3.1). First, we surveyed candidate islands to make sure there were no resident *A. sagrei* (as informed initially by D.A.S.'s decades of experience conducting fieldwork in this area, and subsequently confirmed during surveys of each island). We chose eight islands to serve as destinations for transplants. We then used the *Polygon Area* function implemented in Google Earth to measure vegetated area of each island (Table S3.1), which is strongly associated with population density of *A. sagrei* (Kolbe et al. 2012). In addition, we used data loggers to record temperatures throughout the day to evaluate micro-climactic variation. This is an important component of habitat quality because even small increases in the magnitude of temperature variation can lead to decreases in physiological performance in tropical ectotherms such as *Anolis* lizards (Logan et al. 2014; Akashi et al. 2016). We placed three data loggers (Onset, Hobo®) on the outer-rim (non-vegetated, rocky shoreline), the boundary of vegetation edge, and the inner, vegetated core of each island. From these data, we calculated indices of vegetation and island quality: 1) the vegetated area of each island scaled for the mean vegetated area of all islands, and 2) the average daily temperature difference recorded across the sampling period on each island scaled for the average daily temperature difference across all islands. Using these scores, we performed hierarchical clustering with Euclidean distances (function *hclust* in R) to cluster islands into high- and low-quality clusters (Fig. 3.2). Thus, high-quality islands (109, 110, 409, 911) had greater vegetated area and more stable temperatures, while low-quality islands (206, 914, 917, 932) had less vegetation and more variable temperatures (i.e., on average, higher maximum and lower minimum temperatures). In addition, the daily maximum temperature in low-quality islands exceeded the heat tolerance limits (CTmax) of *A. sagrei* (ranging from 38 to 43 ºC; Corn 1971; Gunderson et al. 2018), suggesting that these habitats represented highly stressful thermal environments.

After identifying and surveying destination islands, we collected a total of 227 *A. sagrei* adults (140♀:87♂) in Staniel Cay, Bahamas during April 2016. We tagged each individual with coloured elastomer (Northwest Marine Technology, Inc., USA). We immediately dissected eight individuals on Staniel Cay to serve as controls, and transplanted the remaining individuals onto experimental islands, with the number of individuals based on

the island vegetation area (0.1 individuals/m^2 of vegetated area) and a sex ratio of approximately 1.6♀:1♂, following the experimental set-up described in Schoener and Schoener (1983). Following at least four days of exposure, we recaptured lizards from each island, and immediately sacrificed and dissected individuals with matched body size (males > 40 mm, females > 38mm in SVL) across treatments to remove livers. We selected liver because it is involved in a number of important processes in animals, including bone metabolism (Sjögren et al. 2002), temperature tolerance (Orczewska et al. 2010), energy storage (Chellappa et al. 1989), and immune function (Wenzel and Piertney 2014). In total, we sampled livers from 36 lizards: eight (4♀ and 4♂) from the source island, 13 (4♀ and 9♂) from high-quality islands, and 15 (8♀ and 7♂) from low-quality islands (Table S3.2). We extracted genomic DNA using phenol:chloroform:isoamyl alcohol (25:24:1), and assessed DNA quality and quantity using Tecan Infinite® 200 NanoQuant and Quant-iT™ PicoGreen® dsDNA Assay Kit (ThermoFisher Scientific, USA).

S3.2. Reduced representation bisulphite sequencing

To measure genome-wide DNA methylation levels, we constructed reduced representation bisulphite sequencing libraries using a previously published protocol (Boyle et al. 2012). Briefly, genomic DNA (200 ng) for each individual was digested using the restriction enzyme *MspI* (New England Biolabs, USA), and ligated to a unique methylated Illumina TruSeq adapter. We targeted fragments of 160-340 bp (including ~120 bp of adapter sequence) using NaCl-PEG diluted SpeedBeads (Rohland and Reich 2012). Libraries were treated with the EpiTect Bisulfite kit (Qiagen, Germany) to convert unmethylated cytosines to thymine after PCR. The resulting 36 libraries were amplified by PCR with Illumina TruSeq primers, pooled in batches of nine, and sequenced in four lanes (100-bp single-end reads, nine samples per lane) within the same flow cell of an Illumina HiSeq 2500 at the McGill University and Genome Quebec Innovation Center. Each sample was sequenced to a mean depth (±SD) 19.53 ± 8.249 of million reads, providing sufficient data for downstream analysis (Gu et al. 2011; Boyle et al. 2012). Three of the 36 samples did not have at least 10 million reads and were excluded from further analysis (samples r8, r65, and r99 in Table S2). Average mapping efficiency was 39.32 ± 0.83% (± SD). We also quantified methylation at non-CpG motifs, and found less than 1% of non-CpG cytosines were methylated, suggesting a highly efficient bisulphite conversion.

S3.3. Read alignment, DMC and DMR calling, and gene ontology analysis

To remove adapter contamination and low-quality reads, we performed read quality checks using FastQC v0.11.5 (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc), and trimmed low-quality and adapter-contaminated sequences using Trim Galore! v0.4.4 (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) with the 'rrbs' option. We then ran Bismark v0.17.0 (Krueger and Andrews 2011) with the Bowtie2 v2.2.8 aligner (Langmead and Salzberg 2012) against a draft *A. sagrei* reference genome with default settings, except for tolerating one non-bisulphite mismatch per read. We only included reads that mapped uniquely to the reference genome in downstream differentially methylated CpG sites (DMCs) and differentially methylated regions (DMRs) analyses.

To characterize the immediate response to environmental change in the founding populations, we analyzed DMCs and DMRs in three comparisons: high-quality islands vs. low-quality islands, high-quality vs. the source island, and low-quality vs. the source island to examine specific epigenetic responses to each island quality. Before running the formal analyses, we tested the effects of sex and island on methylation variation among islands within each island category, and among island categories; however, neither of the two factors was significantly associated with the first principal component (explaining >95% the variation in the methylation values) calculated from the percent methylation matrix (Akalin et al. 2012). In addition, we observed very high positive correlation (ranging from 0.96 to 0.97) between samples from the same island, and between samples from the same island category regardless of sex, suggesting that sex-specific or island differences in methylation are low. We thus excluded sex and island as covariates in the formal analyses, and treat all individuals from the same island category as replicates.

We ran the formal analyses by first identifying individual DMCs using the R package methylKit v1.4.1 (Akalin et al. 2012). Read coverage was normalized between samples, using median coverage across all samples as the scaling factor. A minimum of five reads in all samples were required at a CpG site for that site to be analyzed (Walker et al. 2015; Wan et al. 2016). Sites that were in the 99.9th percentile of coverage were also removed from the analysis to account for potential PCR bias. We used default parameters, with a correction for overdispersion, and a minimum required average methylation difference of 25% between lizards from different island categories. Significantly differentially methylated CpG sites were those with a multiple test corrected *P*-value (*Q*-value as per Storey and Tibshirani (2003)) smaller than 0.01. To determine whether the number of DMCs identified in each of the three island type comparisons was greater than expected by chance, we constructed a null distribution using permutations. To build the distribution, we applied the permutation on the

normalized read coverage per site of each individual, and randomly permuted the island type (source, high-, or low-quality islands) between individuals while maintaining the sample ID of each individual. Each permutated data set was analysed by the same steps as the actual data, and repeated 1000 times. We then calculated the probability of observing the number of DMCs given this distribution (Figure S3.2).

We next determined DMRs using the R package eDMR v0.6.4.1 with default parameters (Li et al. 2013). To be considered significant, a DMR needed to contain at least three CpG sites within an algorithm-specified genomic distance, with at least one classified as a DMC (Q-value < 0.01), and an absolute mean methylation difference greater than 20% when comparing lizards from different island categories. We visualized the differential methylation patterns across individuals using the R package pheatmap v1.0.8 (https://cran.r-project.org/web/packages/pheatmap/index.html), and obtained clustering of samples and DMRs in heatmaps with the "complete" clustering method on Euclidian distances. We clustered hyper- and hypomethylated DMRs using the relative percent DMA methylation, which is the normalized percent DNA methylation scaled for each DMR's percent DNA methylation (median per cent methylation as 0). We also clustered individual lizards based on overall methylation patterns across DMRs.

Because the *A. sagrei* genome is not fully annotated, we used BLASTx with default settings against the NCBI non-redundant database to identify DMR-associated genes, followed by functional category assignment, and GO term mapping implemented in Blast2GO v4.1 (Conesa et al. 2005; Götz et al. 2008).

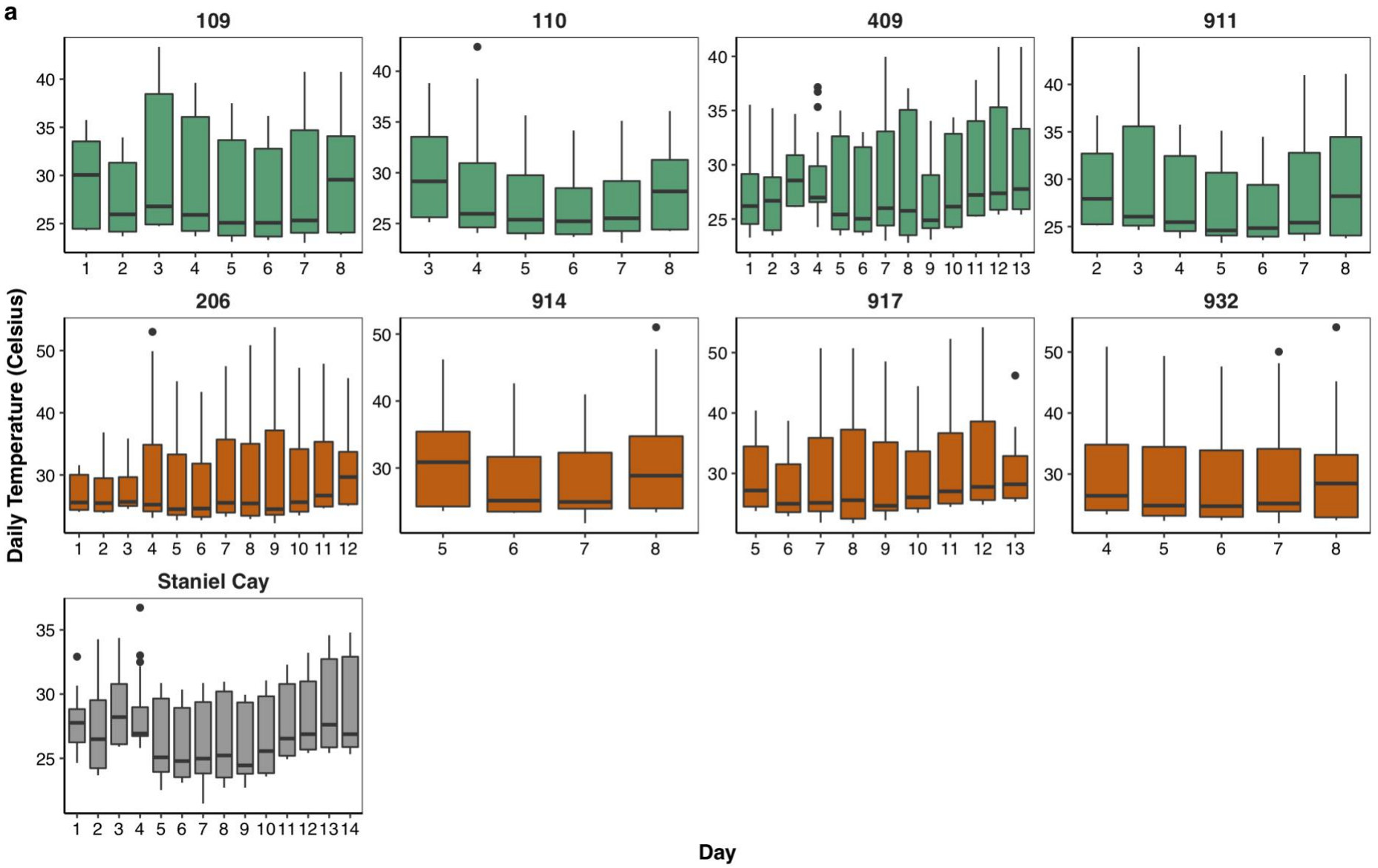
Table S3.1 Island included in this study

Island	Vegetated area (m^2)	Mean temperature difference (mean ± SD ºC)	Latitude	Longitude
Staniel Cay	1,385,637	8.57 ± 1.19	24.170	-76.439
911	216	13.94 ± 4.56	24.240	-76.503
409	262	12.30 ± 3.49	24.204	-76.464
110	324	12.96 ± 3.59	24.221	-76.480
109	510	14.81 ± 2.95	24.221	-76.485
917	40	23.58 ± 6.97	24.192	-76.481
206	112	21.32 ± 7.81	24.239	-76.492
932	404	27.89 ± 2.85	24.209	-76.486
914	484	22.21 ± 5.23	24.188	-76.488

Table S3.2 Sample information, read counts, and alignments to the *Anolis sagrei* genome

Sample ID	Sex	Island ID	Island	Unfiltered	Filtered	Uniquely mapped	No alignment	# of CpG after alignment
r3	M	Staniel Cay	Source	34,134,693	33,653,687	13,163,596	4,290,427	56,112,152
r4	M	Staniel Cay	Source	29,551,229	29,027,078	11,590,576	11,160,953	49,307,144
r5	F	Staniel Cay	Source	25,770,053	25,363,011	25,363,011	6,917,961	43,636,391
r6	M	Staniel Cay	Source	25,638,711	25,280,178	9,970,413	6,562,937	41,685,919
r7	M	Staniel Cay	Source	19,919,106	19,611,297	7,826,139	7,097,002	32,382,105
r8	M	Staniel Cay	Source	6,507,422	6,399,019	2,505,073	4,084,526	10,396,186
r9	F	Staniel Cay	Source	13,086,324	12,866,302	5,196,375	6,473,361	21,869,120
r10	F	Staniel Cay	Source	14,680,173	14,366,912	5,569,958	4,507,790	23,398,960
r69	F	911	High-quality	19,978,644	19,574,592	7,859,676	4,095,948	33,590,220
r78	F	911	High-quality	11,751,125	11,411,379	4,564,251	5,478,159	19,260,486
r83	M	911	High-quality	11,686,700	11,371,757	4,426,764	7,942,101	18,651,688
r89	M	911	High-quality	22,921,750	22,528,158	8,822,353	9,319,465	36,462,630
r99	M	911	High-quality	10,318,037	9,307,290	3,553,175	8,102,600	15,374,155
r65	M	409	High-quality	9,181,546	8,776,142	3,347,764	13,630,337	14,240,167
r77	M	409	High-quality	11,940,413	11,383,395	4,380,245	5,790,012	18,884,387
r84	F	409	High-quality	13,967,421	13,677,589	5,457,908	4,405,845	23,021,644
r94	F	409	High-quality	37,221,166	36,561,781	14,794,279	5,737,017	61,549,447
r29	M	110	High-quality	21,922,028	21,424,992	8,231,290	4,193,951	35,354,196
r38	M	110	High-quality	20,960,449	20,090,105	7,610,055	7,205,812	32,596,884
r91	M	110	High-quality	22,214,839	21,880,050	8,667,656	5,375,505	35,783,817
r93	M	110	High-quality	12,834,423	12,496,579	4,799,023	4,896,303	20,216,893
r24	M	109	High-quality	20,101,770	19,775,248	7,828,511	4,290,427	32,801,071
r33	F	109	High -quality	14,349,647	14,124,219	5,696,225	11,160,953	24,117,437
r62	M	109	High -quality	12,721,590	12,472,141	4,889,854	6,917,961	20,324,481
r37	F	917	Low-quality	17,736,633	17,392,643	6,931,428	6,562,937	29,639,122
r61	F	917	Low-quality	25,996,968	25,002,096	9,904,535	7,097,002	42,645,037
r71	M	206	Low-quality	28,790,032	28,136,693	10,718,684	4,084,526	45,578,480
r73	F	206	Low-quality	26,095,786	25,616,625	10,263,111	6,473,361	43,556,606
r80	F	206	Low-quality	44,053,145	43,386,323	17,493,554	4,507,790	73,760,900

r45	F	932	Low-quality	18,456,712	18,065,750	7,158,470	4,095,948	30,057,041
r75	M	932	Low-quality	13,378,047	12,945,942	4,988,842	5,478,159	21,277,858
r79	F	932	Low-quality	18,123,185	17,740,206	7,004,729	7,942,101	29,029,994
r85	M	932	Low-quality	13,124,700	12,404,641	4,638,218	9,319,465	19,742,913
r32	M	914	Low-quality	22,526,493	22,194,804	8,693,833	8,102,600	36,464,442
r87	M	914	Low-quality	16,149,180	15,757,569	5,968,036	13,630,337	24,828,275
r95	F	914	Low-quality	15,240,914	14,909,457	5,848,996	5,790,012	24,375,846


a
109
110
409
911
206
914
917
932
Staniel Cay
Daily Temperature (Celsius)
Day

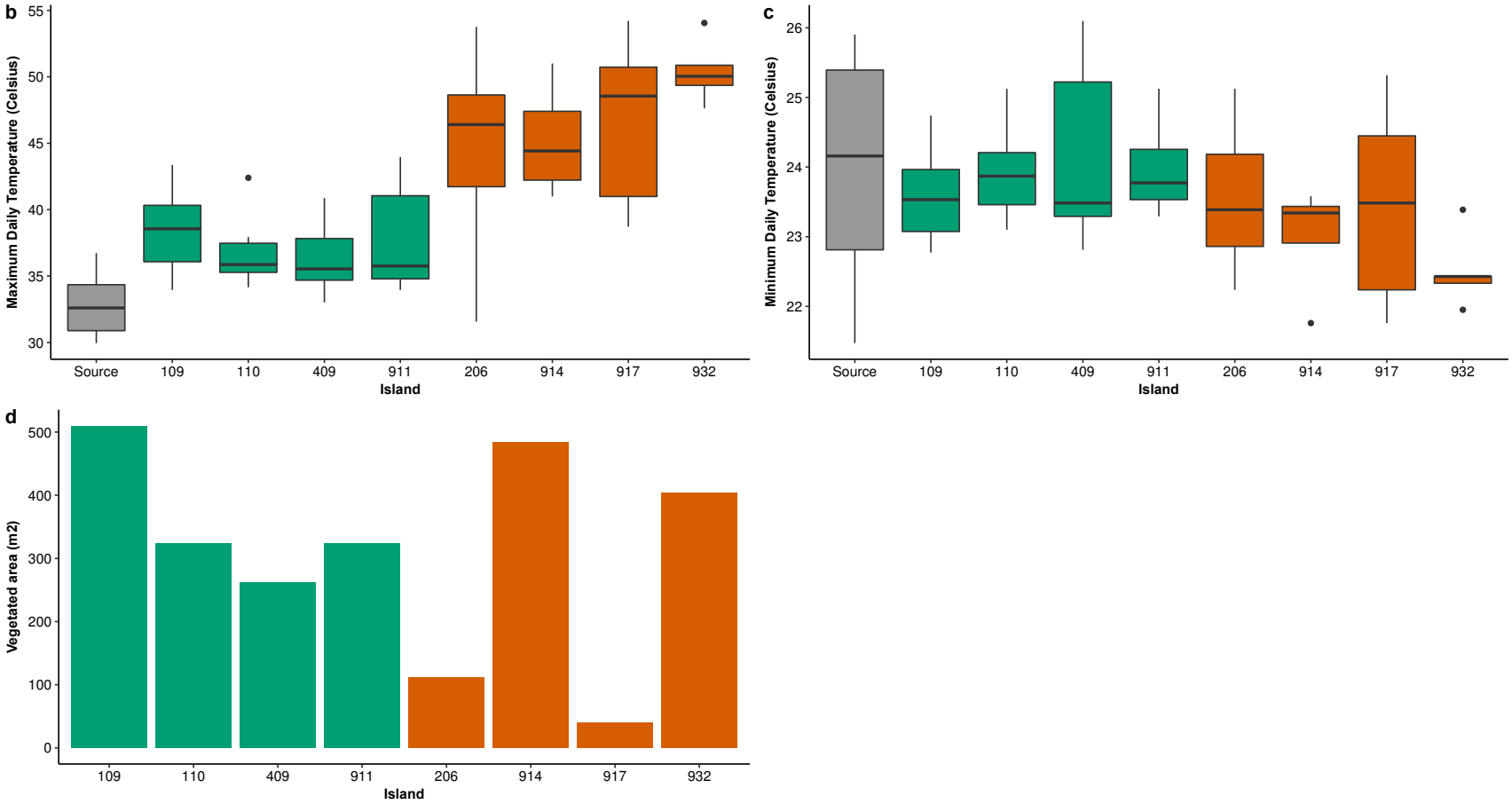


Fig. S3.1 Environmental conditions of the source and treatment islands. High-quality islands are shown in green, low-quality islands are shown in orange, the source island is shown in grey. a) Daily diel temperature, b) maximum daily temperature, c) minimum daily temperature, and d) vegetated area of each island. Daily values for maximum and minimum daily temperatures were averaged from the date that we transplanted lizards to each island to the date that we recaptured them. The black lines in each of the boxes represent the median temperature, the top and

bottom of the boxes represent the 25th and 75th percentile range of temperatures, temperature values that are outside the 1.5 times the inter quantile range (distance between the 25th and 75th percentile range) are represented by dots.

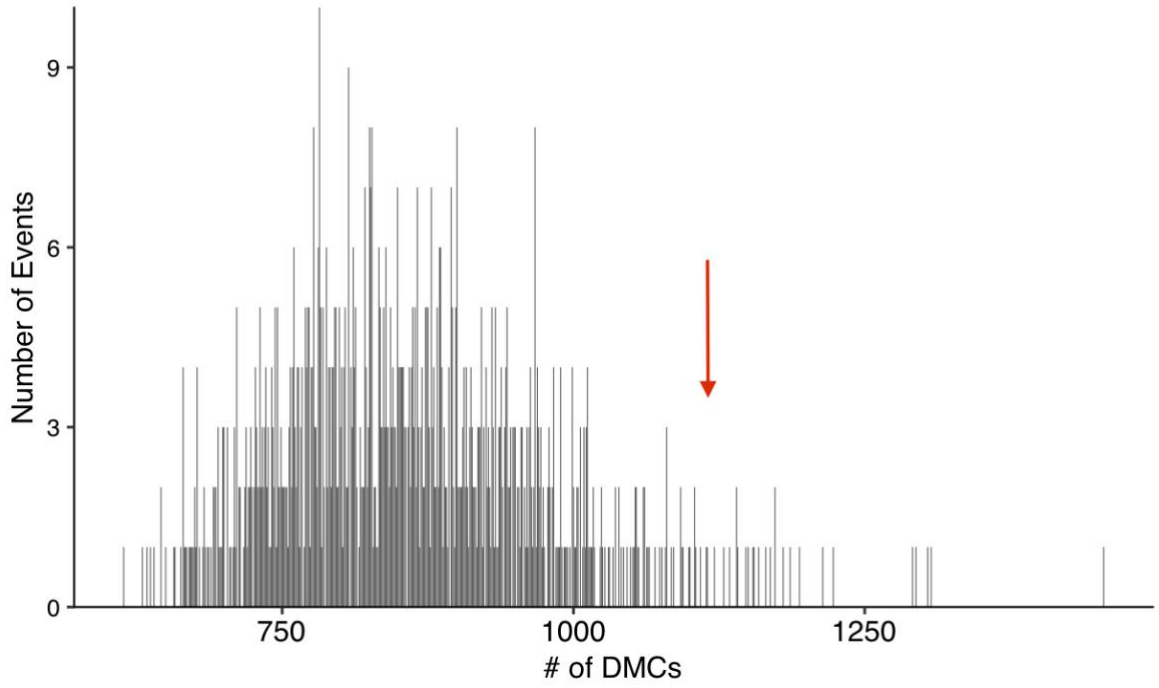


Fig. S3.2 Event histograms of 1000 randomly-generated datasets showing the probability of having the observed number of differentially methylated CpG sites (DMCs) in low-quality vs. source island. Red arrows show the observed number of DMCs.

Supplement text references

Akalin A, Garrett-Bakelman F, Kormaksson M, Busuttil J, Zhang L, Khrebtukova I, Milne T, Huang Y, Biswas D, Hess J, et al. 2012. Base-pair resolution DNA methylation sequencing reveals profoundly divergent epigenetic landscapes in acute myeloid leukemia. *PLoS Genet.* 8:e1002781.

Akashi HD, Cádiz Díaz A, Shigenobu S, Makino T, Kawata M. 2016. Differentially expressed genes associated with adaptation to different thermal environments in three sympatric Cuban *Anolis* lizards. *Mol Ecol.* 25:2273-2285.

Boyle P, Clement K, Gu H, Smith ZD, Ziller M, Fostel JL, Holmes L, Meldrim J, Kelley F, Gnirke A, et al. 2012. Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. *Genome Biol.* 13:R92.

Chellappa S, Huntingford FA, Strang RHC, Thomson RY. 1989. Annual variation in energy reserves in male three-spined stickleback, *Gasterosteus aculeatm* L. (Pisces, Gasterosteidae). *J Fish Biol.* 35:275-286.

Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674-3676.

Corn MJ. 1971. Upper thermal limits and thermal preferenda for three sympatric species of *Anolis*. *J Herpetol.* 5:17-21.

Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 36:3420-3435.

Gu H, Smith ZD, Bock C, Boyle P, Gnirke A, Meissner A. 2011. Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat Protoc.* 6:468-481.

Gunderson AR, Mahler DL, Leal M. 2018. Thermal niche evolution across replicated *Anolis* lizard adaptive radiations. *P R Soc B* 285.

Kolbe JJ, Leal M, Schoener TW, Spiller DA, Losos JB. 2012. Founder effects persist despite adaptive differentiation: a field experiment with lizards. *Science* 335:1086-1089.

Krueger F, Andrews S. 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 27:1571-1572.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Meth.* 9:357-359.

Li S, Garrett-Bakelman FE, Akalin A, Zumbo P, Levine R, To BL, Lewis ID, Brown AL, D'Andrea RJ, Melnick A, et al. (Li2013 co-authors). 2013. An optimized algorithm for detecting and annotating regional differential methylation. *BMC Bioinformatics* 14:1-9.
Logan ML, Cox RM, Calsbeek R. 2014. Natural selection on thermal performance in a novel thermal environment. *Proc Natl Acad Sci U S A.* 111:14165-14169.
Orczewska JI, Hartleben G, Brien KM. 2010. The molecular basis of aerobic metabolic remodeling differs between oxidative muscle and liver of threespine sticklebacks in response to cold acclimation. *Am J Physiol Regul Integr Comp Physiol.* 299:R352.
Rohland N, Reich D. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Res.* 22:939-946.
Schoener TW, Schoener A. 1983. Distribution of vertebrates on some very small islands. II. Patterns in species number. *J Anim Ecol.* 52:237-262.
Sjögren K, Sheng M, Movérare S, Liu J-L, Wallenius K, Törnell J, Isaksson O, Jansson J-O, Mohan S, Ohlsson C. 2002. Effects of liver-derived insulin-like growth factor I on bone metabolism in mice. *J Bone Miner Res.* 17:1977-1987.
Storey JD, Tibshirani R. 2003. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A.* 100:9440-9445.
Walker DL, Bhagwate AV, Baheti S, Smalley RL, Hilker CA, Sun Z, Cunningham JM. 2015. DNA methylation profiling: comparison of genome-wide sequencing methods and the Infinium Human Methylation 450 Bead Chip. *Epigenomics* 7:1287-1302.
Wan ZY, Xia JH, Lin G, Wang L, Lin VCL, Yue GH. 2016. Genome-wide methylation analysis identified sexually dimorphic methylated regions in hybrid tilapia. *Sci Rep.* 6:35903.
Wenzel MA, Piertney SB. 2014. Fine-scale population epigenetic structure in relation to gastrointestinal parasite load in red grouse (*Lagopus lagopus scotica*). *Mol Ecol.* 23:4256-4273.

CONNECTING STATEMENT

In Chapter 3, I analysed epigenetic variation in natural lizard populations transplanted to habitats with distinct quality, and found that the level of methylation variation was depened on environmental shift between newly colonized islands and source island. Several significantly differentially methylated regions mapped to genes annotated with functions that have relevance to both short-term phenotypic variation, and potentially long-term evolutionary change. These results suggest that DNA methylation can facilitate animal responses to real-world levels of environmental complexity. Taken together, the results from Chapters 2 and 3 suggested that environmental change can induce changes in patterns of genome-wide DNA methylation, which may facilitate responses to environmental stressors in natural animal populations. However, the proportion of DNA methylation that is altered by the environment is only a small proportion of the whole genome DNA methylation profile, and as yet only a few studies have been conducted to analyse the function of constitutive DNA methylation that is stable through the life time of non-model animals. In addition, it has been suggested that heritable DNA methylation is rare in mammals, and the heritability of DNA methylation is even less studied in other vertebrates. Thus, the role of heritable DNA methylation variation in ecological and evolutionary processes is unclear.

To add to our understanding of relative proportions and roles of constitutive and heritable methylation, in Chapter 4, I analyse DNA methylation patterns across generations in threespine stickleback fish, and explore the effects of constitutive methylation on development, as well as the implications of heritable methylation for evolutionary processes.

CHAPTER 4: Heritable DNA methylation in threespine stickleback (*Gasterosteus aculeatus*)

Authors: JUNTAO HU,* SARA J. SMITH,† TEGAN BARRY,† SEAN M. ROGERS,† HEATHER A. JAMNICZKY§ and ROWAN D. H. BARRETT*

Initials: J.H., S.J.S., T.B., S.M.R, H.A.J. and R.D.H.B.

Keywords: Constitutive DNA methylation, reduced representation bisulphite sequencing, transcriptional regulation, *Gasterosteus aculeatus*

*Redpath Museum and Department of Biology, McGill University, 859 Sherbrooke Street West, Montreal, Quebec, Canada H3A 0C4
†Department of Biological Sciences, University of Calgary, 2500 University Dr NW, Calgary, Alberta, Canada T2N 1N4
§Department of Cell Biology and Anatomy, McCaig Institute for Bone and Joint Health, University of Calgary, Calgary, Alberta, Canada T2N 1N4

Correspondence: Juntao Hu, Fax: +1 514 398 3185; E-mail: juntao.hu@mail.mcgill.ca

Abstract

Epigenetic mechanisms underlying phenotypic change are hypothesized to contribute to population persistence and adaptation in the face of environmental change. DNA methylation is the most well-studied type of epigenetic modification regulating gene expression. While much research has demonstrated the proportion of methylation altered by the environment, the distribution and function of constitutive methylation in the genomes of natural populations remain elusive. In addition, although the evolutionary potential of methylation is partially related to its transgenerational stability, few studies have explored heritable methylation in natural populations of animals. Here, we use natural populations of threespine stickleback (*Gasterosteus aculeatus*) to characterize general methylation patterns across three generations and explore the functional role of constitutive methylation in gene expression regulation. We quantitatively measured genome-wide DNA methylation in fin tissue using reduced representation bisulphite sequencing of F1 and F2 siblings produced from crosses between ancestral marine and derived freshwater populations. We describe a clear inter-generational methylation pattern that is likely to be shaped by heritable genetic variation. The identified constitutive hypo- and hypermethylated CpG sites in genetically diverse F2 fish were enriched within promoters and gene bodies, respectively, with many CpG sites mapping to genes encoding proteins relevant to fin development and regeneration. We also identified 113 methylated cytosines that showed a highly consistent methylation pattern across generations, and found that genetic variance explained an average of 26% of the methylation variance. These findings demonstrate the importance of constitutive methylation as an epigenetic regulatory mechanism, and improve our understanding of the heritable basis of population epigenomic variation.

Introduction

DNA methylation is a chemical modification to DNA that typically occurs at cytosines within CpG dinucleotides in animals (Suzuki and Bird 2008). DNA methylation is implicated in a number of biological roles, including gene expression regulation and cell-fate decision (Jones 2012). The functions of DNA methylation depend on the genomic context where these modifications occur. For example, methylation at promoters and enhancers is often associated with gene repression, whereas methylation at gene bodies is likely to result in alternative splicing (Maunakea et al. 2010; Jones 2012). While some DNA methylation variation can be plastically altered by the environment (Feil and Fraga 2012), most genomic methylation patterns are somatically static across tissues throughout the lifespan of an organism. This ontogenetically stable methylation typically overlaps with the promoters of housekeeping or development-related genes when hypomethylated (Smith and Meissner 2013), and overlaps with the bodies of actively transcribed genes when hypermethylated (Jones 2012). Considering that physiological activities are usually modulated at the level of gene expression (López-Maury et al. 2008), investigating the locations of constitutively methylated regions across the genome will help elucidate their role in regulating genes during animal development.

Experimental studies investigating heritable DNA methylation and its role in adaptive evolution are in their initial stages (Verhoeven et al. 2016; Hu and Barrett 2017). Although it is clear that some DNA methylation can be inherited across generations (Jablonka and Raz 2009; Daxinger and Whitelaw 2012; Lim and Brunet 2013; Heard and Martienssen 2014), current studies have mainly used isogenic lab animals exposed to external changes under laboratory conditions (but see Nätt et al. 2012; Weyrich et al. 2016). However, such heritable methylation variation can be reset once the environmental cues that initially induced the variation have disappeared, which can preclude identification of the core heritable methylation essential to animal development regardless of environmental change. Furthermore, by tightly controlling genetic background and the environmental stimulus of methylation responses, isogenic lines are not ideal for informing the constitutive methylation patterns occurring in more genetically heterogeneous natural populations.

To study the role of constitutive methylation in animal development, and to explore heritable epigenetic modifications under stable environments, we used threespine stickleback (*Gasterosteus aculeatus*) (Fig. 4.1), an abundant species in both marine and freshwater environments in the Northern Hemisphere. Since the end of the last ice age, several marine

populations colonized freshwater lake habitats that became uplifted and landlocked, resulting in replicate freshwater populations that show repeated evolution of a suite of locally adapted traits (Bell and Foster 1994). The adaptive divergence between freshwater populations and their marine ancestors makes this system an excellent and tractable system for the study of the ecology and genetic architecture of adaptation (Jones et al. 2012). In the last decade, a variety of genetic and genomic resources have been developed for this species (Baird et al. 2008; Hohenlohe et al. 2010; Jones et al. 2012; Huang et al. 2016; Gibbons et al. 2017; Peichel and Marques 2017). In addition, genome-wide methylation variation between certain marine and freshwater populations (Smith et al. 2015) and between males and females (Metzger and Schulte 2018) have been characterized, and methylation responses to environmental change have been demonstrated (Artemov et al. 2017; Metzger and Schulte 2017, 2018). Particularly, Metzger and Schulte (2017, 2018) analyzed within-generational methylation changes in stickleback populations reared at different temperatures or salinities, and found that both temperature and salinity can induce methylation variation that may facilitate the accumulation of epigenetic variation across generations. However, the role of constitutive and heritable methylation in this species remains unclear. Characterizing such variation and its regulation can be assessed under an experimental design with controlled crosses towards elucidating the impact on developmental and evolutionary processes in this system.

Recent studies have suggested that DNA methylation variation is largely controlled by genetic variation (Orozco et al. 2015; Taudt et al. 2016), and that gene expression can be actively regulated by epigenetic modifications in fish (Baerwald et al. 2015; Kratochwil and Meyer 2015). Furthermore, unlike mammals, DNA methylation in fish is not fully reprogrammed, and can be inherited to some extent (Jiang et al. 2013; Potok et al. 2013). Thus, we hypothesize that: (1) Genetic variation is a strong component of inter-individual variation in DNA methylation profiles. (2) Under stable environmental conditions, constitutive methylation regions should map to genes relevant to the normal function and development of the sampled tissue, with constitutive hypomethylation and hypermethylation occurring more frequently in promoters and gene bodies, respectively, and (3) Some DNA methylation will be heritable across generations, and should map to genes encoding proteins essential for cell structure and development. Testing these hypotheses will help to provide a baseline for understanding the heritability of constitutive methylation variation in natural populations.

Results

General methylation patterns across generations

To identify general methylation patterns across generations, we performed a principal component analysis (PCA) on the methylation levels of filtered CpG sites represented in all samples. In total, we identified 54,044 CpG sites that passed the filtering step. The PCA clearly separated the samples by the sire (HL or KL) of F1 populations along the first principal component, which accounted for 12.7% of the variance observed in the data set (Fig. 4.2). Within the same marine-freshwater hybrid line, the second principal component (variance explained: 5.1%) clearly separated parental and F2 samples, with F1 samples filling in the intermediate space between parental and F2 samples.

Constitutive methylation

To identify constitutively methylated sites, we first filtered the cytosines with methylation calls to include only those with less than 10% missing data within HL_F2 or KL_F2 samples (Gugger et al. 2016), and then identified the shared hypomethylated (average DNA methylation levels < 0.1) or hypermethylated sites (average DNA methylation levels > 0.9) between HL_F2 and KL_F2 individuals. The constitutively methylated CpG sites displayed significantly more hypo- than hypermethylation (122,664 hypomethylated vs. 72,609 hypermethylated CpG sites; $\chi^2 = 12{,}831$, df =1, $P < 0.001$). Both hypo- and hypermethylated sites were distributed broadly across the genome, with no apparent clustering on specific chromosomes or chromosomal regions (Fig. 4.3). When analyzing the distribution of constitutively methylated sites around TSSs, we found a clear peak within our defined promoter regions for hypomethylated sites (Fig. 4.4a). In contrast, no peak was found for hypermethylated sites, and more hypermethylated sites were found within the 5kb downstream of TSSs (Fig. 4.4b). In addition, when analyzing genomic context, we found significantly more constitutively hypomethylated sites within promoters (*G*-test; $P < 0.001$), and fewer within gene bodies (*G*-test; $P < 0.001$) and intergenic regions (*G*-test; $P < 0.001$; Fig. 4.4c) when compared to the distribution of CpG sites across the genome. At constitutively hypermethylated sites, we found significantly more sites within gene bodies (*G*-test; $P < 0.001$) and fewer within intergenic regions than expected by chance (*G*-test; $P < 0.001$), with no significant enrichment of promoters (*G*-test; $P = 0.52$; Fig. 4.4d). These results were consistent with our predictions that promoters and gene bodies will be hypomethylated and hypermethylated, respectively.

We found that genes with promoters overlapping with hypomethylated cytosines had more significantly over-represented (false discovery rate < 0.05) GO terms than those that had gene bodies overlapping with hypermethylated cytosines (43 for hypomethylation and 25 for hypermethylation; Table S4.1). Similar to Kratochwil and Meyer (2015), we identified genes involved in processes expected to be over-represented in fin tissue, such as metabolic processes, animal organ morphogenesis, and muscle fiber development. To examine whether genes overlapping with constitutively methylated CpG sites include housekeeping genes or genes that are of importance for fin-specific cellular processes, we performed a review of recent literature to acquire a list of housekeeping genes (Hibbeler et al. 2008), and genes involved in the fin development (Akimenko et al. 1995; Offen et al. 2009), regeneration (Katogi et al. 2004; Kratochwil and Meyer 2015) and structural maintenance (van Eeden et al. 1996) (Table 4.1). We identified several housekeeping genes, with some having both hypomethylated promoters and hypermethylated gene bodies. In addition, among all genes identified involving in fin function, more genes exhibited hypomethylation in their promoters, and many of these genes code for transcription factors (e.g., *dlx3b*, *dlx4*), are involved in signalling pathways (e.g., *sufu*, *rar-γ*), and maintain tissue integrity (e.g., *frem2*) (Table 4.1).

Heritable methylation

To identify heritable methylation, we first analyzed differential methylation between HL and KL hybrid lines within the F1 and F2 generations, and then compared the shared differential methylation patterns between the two generations. We identified 1,198 and 520 DMCs between HL_F1 and KL_F1, and HL_F2 and KL_F2, respectively, with 133 of the DMCs shared in both generations. Our permutation analysis indicated that this number of shared DMCs was significantly greater than expected by chance ($P < 0.001$). The shared DMCs showed a highly consistent pattern across generations (Fig. 4.5a), suggesting the existence of heritable methylation. Shared DMCs displayed significantly more hypo- than hypermethylation (105 hypomethylated and 28 hypermethylated DMCs when comparing HL to KL fish; $\chi^2 = 44.579$, df =1, $P < 0.001$). And, while we found significantly more DMCs within exons (*G*-test; $P < 0.001$), and fewer within introns (*G*-test; $P < 0.001$) than expected by chance, DMCs were not significantly enriched in promoters (*G*-test; $P = 0.06$) or intergenic regions (*G*-test; $P = 0.07$; Fig. 4.5b). We annotated the 133 DMCs to 32 genes (Table S4.2); however, we found no significant GO term enrichment (false discovery rate < 0.05). The top GO terms for these genes were protein glycosylation (GO:0006486),

glycoprotein metabolic process (GO:0009100), glycoprotein biosynthetic process (GO:0009101) and macromolecule glycosylation (GO:0043413) in Biological Process, Golgi subcompartment (GO:0098791) in Cellular Component, and molecular function regulator (GO:0098772) in Molecular Function.

To test the associations between SNP variation and methylation levels, we first identified 92,984 SNPs, including 408 SNPs with less than 10% missing data across all F1 and F2 individuals after filtering. Among the 408 SNPs, six SNPs were revealed by the linear mixed model to have highly significant associations with the methylation values of F1 and F2 individuals ($Q < 0.01$). None of these six SNPs were located near the 133 shared DMCs when analyzing their genomic locations (mean distance = 3342234 bp; distance range: 113927-7483393 bp). Finally, we calculated the kinship matrix using the 408 post-filtering SNPs, and estimated the narrow sense heritability for CpG methylation levels was on average 26%.

Discussion

The role of DNA methylation in fundamental ecological and evolutionary processes has received increased attention in recent years (Verhoeven et al. 2016; Hu and Barrett 2017). However, little is known about the role of constitutive methylation during development, and to what extent variation in DNA methylation is transmitted across generations in natural animal populations. We used a quantitative, high-resolution technique (RRBS) to measure DNA methylation in sticklebacks across three generations. We found that epigenetic variation was associated with genetic variation to some extent, with a narrow sense heritability of 26%. This value is consistent with recent epigenome-wide association studies that have found that genetic variation can explain an average of 7-34% of methylation variation in animals (Orozco et al. 2015; Taudt et al. 2016). We also found distinct distribution patterns between constitutively hypo- and hypermethylated CpG sites: hypomethylated sites occur more frequently around TSSs and are predominantly located within promoters, whereas hypermethylated sites are more likely to be found in gene bodies. In addition, many of the constitutively methylated CpG sites overlapped with genes that are expected to be actively transcribed in caudal fin tissue, with gene ontology analysis showing an over-representation of processes expected to be enriched in fin tissue (e.g., endothelial cell proliferation). Finally, we identified a number of DMCs showing a highly consistent methylation pattern across generations, with many of the CpG sites overlapping with genes encoding proteins that are fundamental for cells. Overall, our study provides the first

investigation of constitutive methylation involved in fin-specific cellular processes, and identified epigenetic variation that is heritable across two generations.

Methylation patterns across generations

Our PCA analysis of general methylation patterns clearly separated the two lines (HL and KL) of marine-freshwater hybrids, and also differentiated the parental, F1, and F2 individuals within the same hybrid line. Many factors can impact methylation variation, including genetic variation, environmental variation, and stochastic epimutation (Shea et al. 2011; Taudt et al. 2016; Hu and Barrett 2017). Because we reared all individuals in a constant environment, and the contribution of stochastic epimutation to DNA methylation variation is typically only observed over a large number of generations (Schmitz et al. 2011), the most likely interpretation of this differentiation is that it is driven by genetic differences between groups. PC1 appears to be associated with genetic variation between the parental HL and KL freshwater populations, whereas PC2 may reflect an increased level of genetic diversity in F2 generations. Despite using a different tissue type, a recent study exploring the epigenetic responses of marine and freshwater sticklebacks exposed to different salinity environments has also suggested that genetic background plays an important role in shaping epigenetic variation (Artemov et al. 2017). Together with this previous work, our results suggest that epigenetic variation is tightly linked to underlying genetic variation.

Linking constitutive methylation to genomic architecture and biological functions

We found a peak distribution of constitutively hypomethylated CpG sites around TSSs, with over half of the hypomethylated sites located within promoters, a proportion significantly greater than expected by chance. This is consistent with previous findings that CpGs at promoters remain predominantly unmethylated in somatic cells (Jones 2012). In contrast, constitutively hypermethylated CpG sites showed a biased distribution to downstream of TSSs. Indeed, this bias is supported by the fact that more than 50% of hypermethylated sites were found within gene bodies, a proportion that is greater than expected by chance. Unlike promoter methylation, gene body methylation is not associated with repression, and is a feature of transcribed genes in vertebrates (Wolf et al. 1984; Meissner et al. 2008; Laurent et al. 2010). Recent studies have also suggested extensive positive correlations between active transcription and gene body methylation (Hellman and Chess 2007; Lister et al. 2009; Feng et al. 2010). Taken together, the large proportion of constitutively methylated CpG sites suggest

that many genes are in an active or poised state in the caudal fin, making it a useful tissue for future ecological and evolutionary studies.

When analyzing the functions of genes that had promoters or gene bodies overlapping with constitutively hypo- or hypermethylated sites, we found several housekeeping genes, as well as genes involved in fin development, regeneration, structure maintenance. We also observed hypomethylation at promoters of many transcription and signalling-related factors (e.g., *hox* and *dlx* genes). This suite of genes can provide cells with basal functions that are essential for cell maintenance, and more importantly, a mechanism that is poised for activation when fish are injured, by equipping the cells with a positional memory. It is likely that these genes are actively transcribed during regenerative processes, where intricate tissues such as the caudal fin can be faithfully restored. Interestingly, we found a significant overlap between the genes identified in our study and those identified in a previously study, which used histone mark trimethylated Histone H3 Lysine 4 (H3K4me3) to find active promoters in fins (Kratochwil and Meyer 2015). This may suggest that instead of isolated epigenetic mechanisms, it is often suites of epigenetic mechanisms that act in concert to regulate animal phenotypes.

In association with similar studies investigating immune response in fish (e.g., Haase et al. 2014; Huang et al. 2016), we found genes with hypomethylated promoters involved in the oxidation-reduction process (GO:0055114). This is likely to be accompanied by upregulation of genes providing antidoxidants, which can limit the negative effects of reactive oxygen species (ROS) caused by fin injury and infection. This mechanism of defence is general, and can combat against both pathogens and macroparasites (Feis et al. 2018). Another significantly enriched GO term is the Notch signaling pathway (GO:0007219). Genes involved in this pathway that have hypermethylated gene bodies can play roles in immune systems by influencing multiple lineage decisions of developing lymphoid and myeloid cells, and modulating T cell-mediated immune responses (Radtke et al. 2010). Taken together, it is evident that a number of genes related to immune response were actively transcribed in fin tissues, which is not surprising given that the fin is a primary site of inital pathogen and parasite entry in fish (e.g., Baerwald 2013).

Heritable epigenetic loci

While very few studies have analyzed transgenerational epigenetics using natural animal populations (e.g., Nätt et al. 2012; Weyrich et al. 2016), examination of heritable methylation

sites is valuable for predicting the likelihood that populations will be able to evolve in response to environmental change (O'Dea et al. 2016; Hu and Barrett 2017). We identified 133 such CpG sites showing consistent methylation differences between HL and KL populations across generations, a number significantly greater than expected by chance alone. These loci are candidates for transgenerational epigenetics, but notably represent only a very small proportion (~0.1%) of the CpG sites retained by RRBS (~1% of all CpG sites in the stickleback genome), suggesting that epigenetic reprogramming leaves few possibilities for inheritance of epigenetic modifications (Heard and Martienssen 2014). While the proportion of heritable CpG sites is low, these candidate loci were involved in physiological processes necessary for cell structure and development, consistent with our predictions (Table S2). For example, many CpG sites were mapped to genes involved in glycosylation or glycoprotein-related processes, which produce cellular glycans that are frequently attached to proteins and lipids and participate in many key biological processes including cell adhesion, molecular trafficking and clearance, receptor activation, signal transduction and endocytosis. These processes are essential in the development and physiological of living organisms (Ohtsubo and Marth 2006). As for the contribution of these heritable CpG sites to potential evolutionary responses, examination of their genomic context showed enrichment within exons, suggesting that heritable DNA methylation may facilitate alternative splicing, and thus contribute to genomic evolution by increasing the number of transcriptional opportunities and phenotypes (Roberts and Gavery 2012). Consequently, in combination with its higher mutation rate than genetic mutation, heritable methylation could accelerate the exploration of phenotypic space, and thereby allow populations to adapt to the changing environments more efficiently (Klironomos et al. 2013; Hu and Barrett 2017).

Our heritable methylation analysis raises the question of whether heritable methylation patterns have a genetic basis. When analyzing the associations between SNPs and methylation patterns, we identified six SNPs that showed significant associations with methylation patterns. None of these SNPs were located near the identified 133 CpG sites that are candidate loci for heritable methylation. Thus, this analysis does not provide clear candidate SNPs for heritable methylation. However, our results do not exclude the possibility that these SNPs act in *trans* on methylation patterns. *Trans*-acting SNPs can explain a modest proportion of methylation patterns in model animals, yet the underlying mechanisms are not well understood (Taudt et al. 2016), reinforcing that further studies exploring the role of *trans*-acting SNPs would be valuable. We also examined the degree to which methylation is controlled by genetics by using methylation levels of individual CpG sites as phenotypes. We

observed an average heritability of 26% for all CpG sites across F1 and F2 generations, suggesting that despite the lack of clear *cis*-acting candidate SNPs for heritable methylation variation, constitutive heritable DNA methylation levels do show a significant genetic basis.

Limitations

An intrinsic problem of *in vivo* studies using next-generation sequencing techniques such as RRBS is the heterogeneity of analyzed tissues. Fin tissues consist of many different cell types including epidermis, osteoblasts, dermal fibroblasts, and vascular endothelium (Tu and Johnson 2011). Therefore, various proportions of different cell types could introduce biases in measures of methylation levels (Kratochwil and Meyer 2015). However, the stringent filtering parameters we used to identify constitutive and heritable methylation may reduce these potential biases by only retaining the high quality reads, and CpG sites that are likely independent of sequence variation across all individuals. In addition, while methylation patterns are tissue-specific, the involvement of fin tissues in various responses to environmental perturbations provides the advantage of testing methylation patterns associated with a wide range of physiological processes (Baerwald 2013; Yoshida et al. 2014; Baerwald et al. 2015; Jesus et al. 2016; Mogi et al. 2017).

It is also important to note that the reduced representation genome sequencing method used here can only cover a small proportion of all possible constitutive and heritable methylation patterns that may exist in stickleback. In addition, while we mainly focused methylation patterns in promoters and gene bodies, other regulatory elements such as enhancers and transposons, although less well annotated in stickleback, are also important drivers of regulatory and phenotypic evolution (Wittkopp and Kalay 2011). Finally, although we identified a number of DMCs that are candidate loci for heritable methylation, RRBS does not permit discrimination between existing C-T DNA polymorphism in the stickleback genome and epigenetic variation. Thus, our results are preliminary, and likely to represent an underestimate of constitutive and heritable methylation modifications. A wider investigation of regulatory elements in combination with genome-wide sequencing of chromatin modifications (e.g., chromatin immunoprecipitation followed by sequencing (ChIP-seq); Park 2009; Furey 2012) and whole-genome resequencing (e.g., Le Luyer et al. 2017) will provide a more comprehensive and precise understanding of the roles that constitutive and heritable epigenetic responses may play in facilitating evolutionary change.

Conclusions

Here, we provide the first insights into patterns of constitutive and heritable methylation in threespine stickleback. Our genome-wide methylation data suggests a small but consistent contribution of heritable, genetically-based epigenetic variation. It also reveals genes and regulatory networks that are likely to be active or poised in fin tissues, and provides candidate loci that are of importance to evolution and development. As predicted, we found constitutively hypo- and hypermethylated loci were predominantly located within promoters and gene bodies, respectively. Candidate loci for heritable methylation only represent a small proportion of all CpG sites, possibly due to germline reprogramming. Our study adds to the few studies using outbred vertebrates to test for heritable epigenetic variation and its genetic basis, and suggests that constitutive methylation can play an important role in regulating gene expression.

Materials and methods

Sampling and husbandry

We collected adult threespine stickleback (*Gasterosteus aculeatus*) from one marine (Bamfield Inlet, BI, 48°49'12.69”N, 125° 8'57.90”W) and two freshwater (Hotel Lake, HL, 49°38'26.94”N, 124° 3'0.69"W, and Klein Lake, KL, 49°43'32.47”N, 123°58'7.83"W) locations in South-western British Columbia, Canada in May 2015 (Fig. 4.1). We transported all fish to our aquatic facility at the University of Calgary, and separated them into population-specific 113 L glass aquaria. We maintained a density of approximately 20 fish per aquarium, salinity of 4-6 ppt, water temperature of 15 ± 2 °C, and a photoperiod of 16 L: 8 D. We kept each aquarium as separate closed systems with its own filter, air pump, water supply, and temperature regulator. We fed all fish *ad libitum* once per day with thawed bloodworms (Hikari Bio-Pure Frozen Bloodworms).

Crossing design

To generate hybrid marine-freshwater F1 families, we collected eggs from marine females, and extracted testes from freshwater males. To generate a cross, we first equally distributed the eggs into a Petri dish containing fresh water. We then euthanized the male using an overdose of eugenol, and removed the testes. We crushed the testes in the Petri dish, with the water activating the released sperm and allowing fertilization. Fertilized eggs were left within the Petri dish for 20 minutes before being suspended in a well-aerated mesh-bottom container

within a 37 L glass aquaria, with an air stone for oxygenation and a sponge filter. In total, we produced two F1 families of HL hybrids (hereafter referred to as HL_F1), and three F1 families of KL hybrids (hereafter referred to as KL_F1). After hatching, the larval fish were reared in the cross-specific 37 L aquaria until reaching approximately 1 cm total length (TL), at which time the families were equally split into 113 L glass aquaria to maintain low densities. The larval fish and fry were fed twice daily with live *Artemia spp*. nauplii. Juvenile stickleback fish were transitioned to a diet of chopped thawed bloodworms at approximately 2 cm TL once per day *ad libitum*, and transitioned to the adult diet of full thawed bloodworms gradually. We sampled fin clips when individuals reached a 3.5 cm TL or more. In addition to the fish we used to generate the F1 crosses, we also sampled extra parental fish from the same population. Fin clips were stored in 70% ethanol in microcentrifuge tubes at room temperature until extraction of genetic material.

To generate F2 families, one male and one female from each family were crossed with the same method used to generate the F1 families. In total, we produced two F2 families of HL hybrids (hereafter referred to as HL_F2), and five F2 families for KL hybrids (hereafter referred to as KL_F2). Fish were raised as described above. We randomly selected fish from one HL_F2 and one KL_F2 family, and sampled caudal fin clips when individual reached approximately 3.5 cm TL. We stored all fin clips as described above. In addition to the fish we used to make the F2 crosses, we also randomly sampled extra F1 fish from all F1 families. In total, we sampled 11 parental fish (six marine females; two HL and three KL freshwater males), 19 F1 fish (7 HL_F1 and 12 KL_F1), and 64 F2 fish (28 HL_F2 and 36 KL_F2). All sampling, crossing, and housing protocols were approved by the University of Calgary Life and Environmental Science Animal Care Committee (AC13-0040 and AC17-0050) following the Canadian Council for Animal Care ethical standards.

Tissue choice

The choice of tissue used for genome-wide mapping of cytosines can influence the interpretation of methylation patterns (Stricker et al. 2017). We conducted our analyses using caudal fin tissue for several reasons. First, fin morphology is associated with repeated adaptation to divergent marine and freshwater environments in sticklebacks (Walker 2008; Jones et al. 2012), suggesting it is likely to have a heritable basis and may thus be linked to heritable DNA methylation. Second, fin tissues are highly regenerative structures (Katogi et al. 2004), and thus many of the structural and phenotypic features must be actively maintained (van Eeden et al. 1996), suggesting that many genes are in an active or poised

state, which increases the likelihood of finding constitutive methylation. Finally, fins can be dissected quickly and consistently, and the cutting of fin tissue usually does not affect survival.

DNA extraction and sex determination

We extracted DNA from caudal fin using phenol:chloroform:isoamyl alcohol (25:24:1), and assessed the quality and quantity using Tecan Infinite® 200 NanoQuant and Quant-iT PicoGreen® dsDNA assay kit (ThermoFisher Scientific). We determined the sex of each fish following Peichel et al. (2004). All procedures were approved by the University of Calgary Life and Environmental Science Animal Care Committee (AC13-0040 and AC17-0050) following the Canadian Council for Animal Care ethical standards.

Reduced representation bisulphite sequencing

To measure genome-wide DNA methylation levels, we used reduced representation bisulphite sequencing (RRBS) (Meissner et al. 2008; Gu et al. 2011), following Boyle et al. (2012) with some minor modifications. For each individual, we created a library from 120 ng of genomic DNA, and ligated the fragments in each library with unique Illumina TruSeq adapters. We targeted fragments of 160-340bp (including ~120bp of adapter sequence) using NaCl-PEG diluted SpeedBeads (Rohland and Reich 2012). We randomly multiplexed 24 libraries into four pools, and treated the pools with sodium bisulphite (EpiTect, Qiagen) following a protocol for formalin-fixed paraffin-embedded samples (Gu et al. 2011). After two rounds of bisulphite treatment to ensure complete conversion of unmethylated cytosines, each pool was amplified with Illumina primers, and loaded in four lanes (100-bp single-end reads) of a Hiseq 2500 at the McGill University and Genome Quebec Innovation Centre. Each sample was sequenced to a mean depth (± SD) of 8.094 ± 2.532 million reads. The average mapping efficiency was 61.4 ± 4.7% (± SD). We quantified methylation at non-CpG motifs, and found less than 1% non-CpG cytosines were methylated, suggesting a highly efficient bisulphite conversion.

Read filtering and mapping

To remove adapter contamination, low-quality bases, and bases artificially introduced during libaray construction, we trimmed reads using Trim Galore! v0.4.4 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), with the 'rrbs'option.

We then used the program Bowtie2 v2.2.9 (Langmead and Salzberg 2012), implemented in Bismark v0.17.0 (Krueger and Andrews 2011) to align trimmed reads for each sample to the stickleback genome (ENSEMBL version 91) with default settings, except for tolerating one non-bisulphite mismatch per read. We only included reads that mapped uniquely to the reference genome, and cytosines that had at least 10x coverage in downstream constitutive and heritable methylation analyses. Only CpG context cytosine methylation was analyzed because CpG methylation is the most common functional methylation in vertebrates (Suzuki and Bird 2008).

General methylation patterns analysis

To identify general methylation patterns across generations, we performed a principal component analysis (PCA) on methylation levels in all samples, using the *prcomp* function in R (R Core Team, 2018, v3.4.3). We ran the analysis by first identifying cytosines that were covered in all samples using the R package methylKit v1.4.1 (Akalin et al. 2012). Read coverage was then normalized between samples, using the median read coverage as the scaling factor. A minimum of ten reads in all samples was required at a CpG site for that site to be analyzed. We calculated the methylation levels by extracting the total amount of methylation-supporting reads, and the total coverage of each CpG site, using the *percMethylation* function in the R package methylKit. Methylation levels at each CpG site were taken as input variables, whereas each point in multidimensional space represented a stickleback individual.

Constitutive methylation analysis

To identify constitutively methylated sites, we first filtered the cytosines with methylation calls to include only those with less than 10% missing data within HL_F2 or KL_F2 samples (Gugger et al. 2016). We only used F2 fish because they represented the most genetically diverse individuals, and thus increased our chances of discovering constitutively methylated sites independent from sequence variation. We defined constitutively hypo- or hypermethylated sites as average DNA methylation levels less than 0.1 or greater than 0.9 (Lam et al. 2012; Lea et al. 2016). We then identified the shared hypo- or hypermethylated sites between HL_F2 and KL_F2 individuals.

To explore the role of constitutive methylation in regulating gene expression, we first identified hypor- or hypermethylated sites located within genes in the *G. aculeatus* genome,

using the R package GenomicRanges v1.28.6 (Lawrence et al. 2013), and analyzed the distribution of cytosines in a 10 kb window around the transcription start sites (TSSs) using the R package ChIPpeakAnno v3.10.2 (Zhu et al. 2010; Zhu 2013). We then identified the position of constitutively methylated cytosines within genomic features (promoter/exon/intron/intergenic) using the R package genomation v1.6.0. We built a null distribution of genomic features based on all CpG sites in the stickleback genome (Metzger and Schulte 2018). We gave precedence to promoters > exons > introns > intergenic regions when features overlapped and defined promoter regions as upstream 1000 bp and downstream 1000 bp from the TSSs (Akalin et al. 2015). Finally, based on the cytosine distribution and genomic feature result, we performed gene ontology (GO) analysis on genes with promoters or gene bodies overlapping with constitutively hypo- or hypermethylated sites, using the R package topGO v2.28.0 (Alexa et al. 2006), based on a Fisher's exact test. The gene pools we compared hypo- or hypremethylated against were cytosines that passed the coverage filtering step, and that had been used in the constitutive methylation analyses (see above). Over-represented GO terms were those with a multiple-test corrected *P*-values (Benjamini-Hochberg's false discovery rate) below 0.05.

Heritable methylation analysis

To identify heritable methylation, we first analyzed differential methylation between HL and KL hybrid lines within either the F1 or F2 generation, and then compared the shared differential methylation patterns between the two generations. Before running the formal analyses, we tested the effects of family and sex on methylation variation within HL_F1 or KL_F1 samples; however, sex and/or family were not significantly associated with the first principal component (explaining > 95% the variation in the methylation values) calculated from the percent methylation matrix, consistent with previous studies suggesting sex as a non-significant factor on methylation variation (Chatterjee et al. 2013; Lea et al. 2016; Hu et al. forthcoming 2018). We thus excluded sex and family as covariates in the formal analyses.

We ran the formal analysis by first identifying individual differentially methylated cytosines (DMCs) within each generation, using the R package methylKit v1.4.1 (Akalin et al. 2012). Read coverage was normalized between samples, using the median read coverage as the scaling factor. A minimum of ten reads in all samples was required at a CpG site for that site to be analyzed. We removed CpG sites that were in the 99.9th percentile of coverage from the analysis to account for potential PCR bias. We used default parameters (false discovery

rate correction *Q*-value < 0.01), with a correction for overdispersion, and a minimum required methylation difference of 25% between HL and KL fish in each generation (Akalin et al. 2012). We then identified shared DMCs between generations using the R package GenomicRanges v1.28.6 (Lawrence et al. 2013). We analyzed the genomic features of shared DMCs as described above. To determine whether more genes were found to be shared DMCs than expected by chance, we constructed a null distribution using permutations. To build the distribution, we applied the permutation on the normalized read coverage per site of each indivdual, and randomly permuted the hybrid lines (HL or KL) between individuals within either the F1 or F2 generation while maintaing the sample ID of each individual. Each permutated data set was analyzed by the same steps on the acutal data, and repeated 1000 times. We then calculated the probability of observing the number of DMCs that were shared between the F1 and F2 generations given this distribution. We visualized differential methylation patterns across individuals, and clustered shared DMCs in heatmaps with the "complete" clustering method on Euclidian distances, using the R package pheatmap v1.0.8 (https://cran.r-project.org/web/packages/pheatmap/index.html). We clustered the shared DMCs using the relative percent DMA methylation, which is the normalized percent DNA methylation scaled for each DMC's percent DNA methylation (median percent methylation as 0) of HL and KL fish in the heatmap. Finally, we analyzed the genomic contexts and GO terms of the shared DMCs to identified their genomic features and functional enrichment.

To determine the genetic basis of methylation patterns across F1 and F2 generations, we used the methylation value of each CpG site of all F1 and F2 individuals for input to Bis-SNP v0.82.2 (Liu et al. 2012) with the default parameters. We then used GATK's VariantFiltration and SelectVariants to restrict variants to diallelic sites, and filter variants based on the following GATK variant annotation cut-offs: QD < 2.0, MQ < 40.0, MQRankSum < -12.5, and ReadPosRankSum < -8.0. We only retained SNPs with less than 10% missing data across all individuals for the downstream anaylsis to minimize false positives. We then used linear mixed models in PyLMM (http://genetics.cs.ucla.edu/pylmm/index.html) to test whether variation at any of the SNPs is significantly associated with methylation patterns in F1 and F2 generations, after adjusting for kinship based on the SNP data. We adjusted *P*-values to *Q*-values using a multiple test as per Storey and Tibshirani (2003).

Finally, we estimated the narrow sense heritability of DNA methylation levels for individual CpG sites, using a linear mixed model approach (Yang et al. 2010; Orozco et al. 2015) implemented in the R package lmmlite (https://github.com/kbroman/lmmlite). We

treated the methylation levels at individual CpG sites of all F1 and F2 individuals as phenotypes, and assumed each phenotype **y** can be modelled as $\mathbf{y} = \mathbf{1_n}\boldsymbol{\mu} + \mathbf{u} + \mathbf{e}$, where the random variable **u** follows a normal distribution centered at zero with variance $\sigma_g^2\mathrm{K}$, and **e** represents an independent noise component with variance σ_e^2. The matrix K is the same kinship matrix as calculated above. For each trait we estimated σ_g^2 and σ_e^2 using the restricted maximum likelihood (REML) approach, and calculated the heritability as $h^2 = \sigma_g^2/(\sigma_g^2 + \sigma_e^2)$. Finally, we calculated the average heritability by taking the mean of heritability values of individual CpG sites.

Acknowledgements

The authors would like to thank Antoine Paccard and Caroline LeBlond for their assistance in sample preparation. This work was supported by China Scholarship Council Fellowship 201406350023 to JH, Queen Elizabeth II Scholarship to SJS and TB, Alberta Innovates to TB, a NSERC RT735278 to SMR, a NSERC DG 418249-2012 to HAJ, and a NSERC Discovery Grant and Canada Research Chair to RDHB.

Data Accessibility

Raw Illumina sequencing reads for the 94 analyzed individuals can be downloaded from the NCBI Short Read Archive (SRA accession: XXXXXXXXX).

Author Contributions

JH and RDHB conceived the study. SJS, TB, HAJ, and SMR sampled and reared fish and collected tissue. JH generated and analyzed sequencing data. JH wrote the manuscript with input from SS, SMR, and RDHB. The authors have no conflicts of interest.

Table 4.1 List of housekeeping genes and genes with implications for fin development and regeneration

Symbol	Gene name	Process involved	Hypo/Hyper	Citation
actb	Beta-actin	Housekeeping gene	Hypo	Hibbeler *et al*. 2011
axin1	Axin 1	Fin regeneration	Hyper	Katogi *et al*. 2004
bmp4	Bone morphogenetic protein 4	Fin regeneration	Hyper	Katogi *et al*. 2004
dkk3a	Dickkopf WNT signaling pathway inhibitor 3a	Fin regeneration	Hypo	Katogi *et al*. 2004
dlx3b	Distal less homeobox protein 3b	Fin regeneration	Hypo	Kratochwil *et al*. 2015
dlx4b	Distal less homeobox protein 4b	Fin regeneration	Hypo & Hyper	Katogi *et al*. 2004
eef1a	Translation elongation factor 1a	Housekeeping gene	Hypo & Hyper	Hibbeler *et al*. 2011
fgfr1	Fibroblast growth factor receptor1	Fin regeneration	Hyper	Katogi *et al*. 2004
frem2	Fras1-related extracellular matrix protein 2	Structure maintenance	Hypo & Hyper	van Eeden *et al*. 1996
g6pd	Glucose-6-phosphate dehydrogenase	Housekeeping gene	Hyper & Hyper	Hibbeler *et al*. 2008
hoxa13b	Homeobox A13b	Fin regeneration	Hypo	Katogi *et al*. 2004
hoxc13a	Homeobox C13a	Fin regeneration	Hypo	Katogi *et al*. 2004
hprt1	Hypoxanthine phosphoribosyltransferase 1	Housekeeping gene	Hypo	Hibbeler *et al*. 2011
junb	Jun B proto-oncogene	Fin regeneration	Hypo	Katogi *et al*. 2004
klf2	Kruppel-like factor 2	Fin development	Hypo	Offen *et al*. 2009
klf5	Kruppel-like factor 5b	Fin regeneration	Hypo & Hyper	Katogi *et al*. 2004
lmo4b	LIM domain only 4b	Fin regeneration	Hypo	Katogi *et al*. 2004
msx2	Msh homeobox 2	Fin regeneration and development	Hypo	Akimenko *et al*. 1995
notch3	Neurogenic locus notch homolog protein 3	Fin regeneration	Hyper	Katogi *et al*. 2004
ppiaa	Peptidylprolyl isomerase A	Housekeeping gene	Hypo & Hyper	Hibbeler *et al*. 2008
rar-γ	Retinoic acid receptor gamma	Fin regeneration	Hypo & Hyper	Katogi *et al*. 2004
sdf2	Stromal cell derived factor 2	Fin regeneration	Hypo	Katogi *et al*. 2004
sfrp1a	Secreted frizzled-related protein 1a	Fin regeneration	Hypo	Katogi *et al*. 2004
smad2	SMAD family member 2	Fin regeneration	Hyper	Katogi *et al*. 2004
smad4	SMAD family member 4	Fin regeneration	Hypo	Katogi *et al*. 2004
sox4	SRY-box 4	Fin regeneration	Hypo	Katogi *et al*. 2004
stat5b	Signal transducer and activator of transcription 5B	Fin regeneration	Hyper	Katogi *et al*. 2004
sufu	Suppressor of fused homolog	Fin regeneration	Hypo	Katogi *et al*. 2004
taf2	TATA-Box binding protein	Housekeeping gene	Hyper	Hibbeler *et al*. 2011

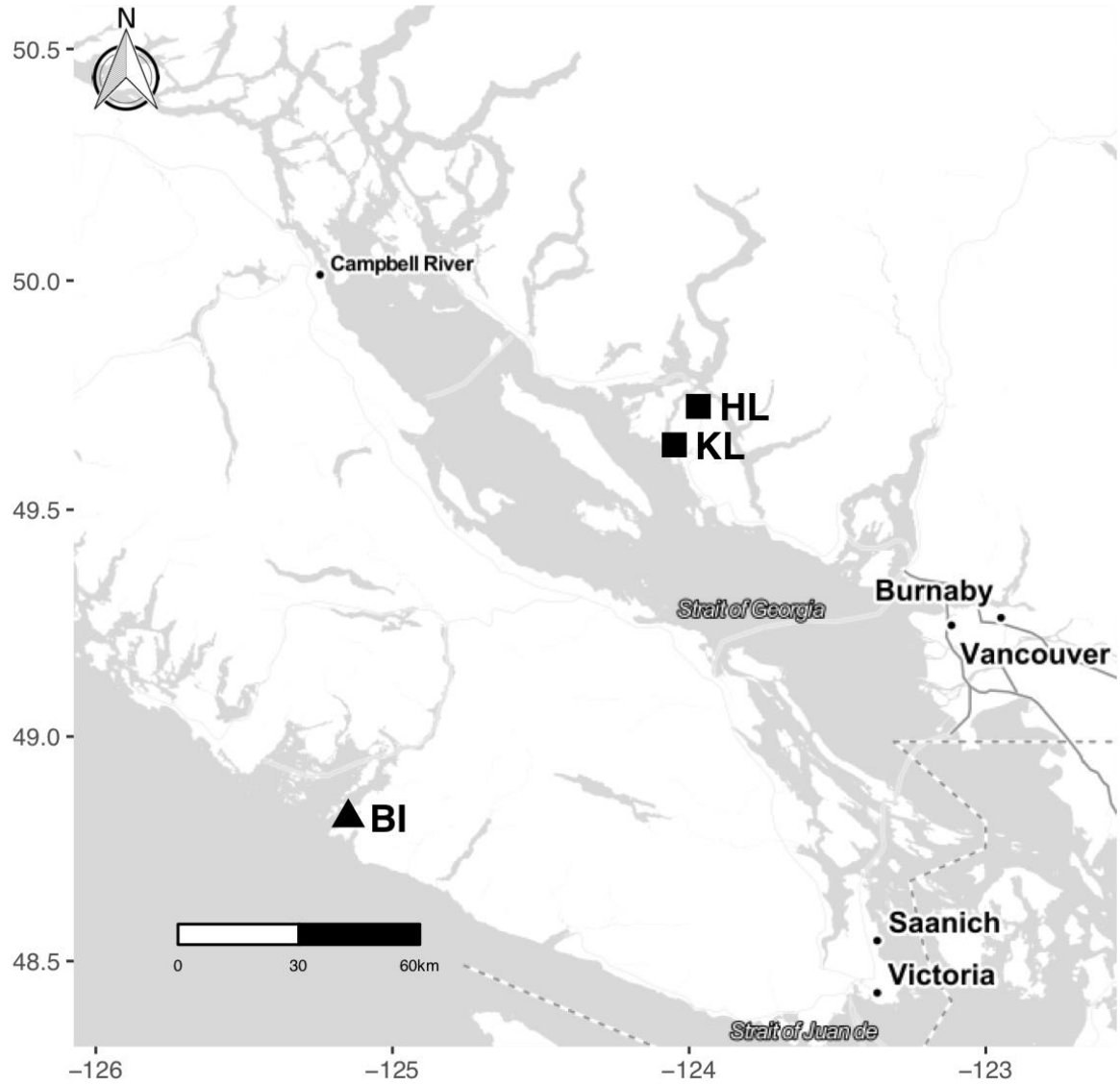


Fig. 4.1 Geographical location of threespine stickleback populations used in this experiment. Triangle indicates the marine sampling site, squares indicate freshwater sampling sites. BI, Bamfield Inlet (marine); HL, Hotel Lake (freshwater); KL, Klein Lake (freshwater).

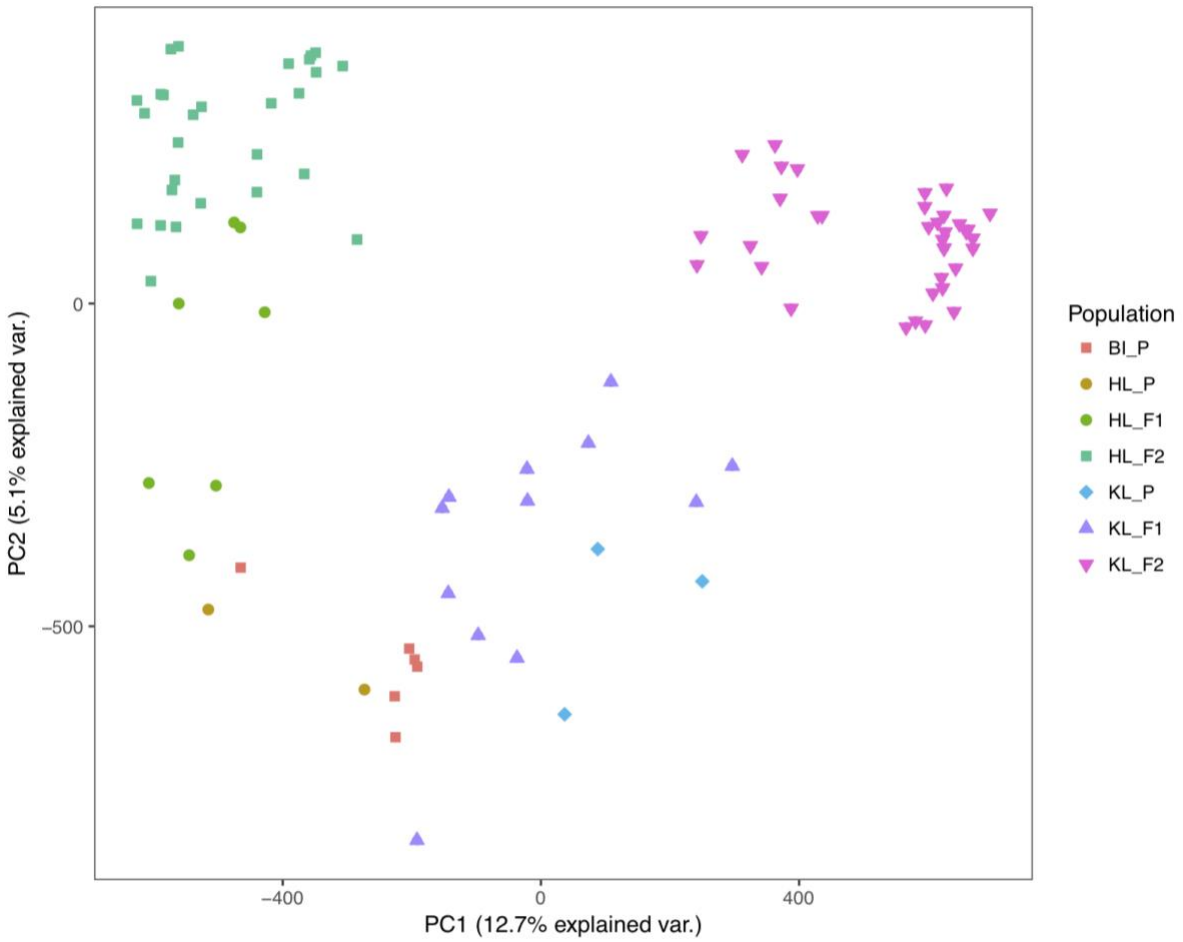


Fig. 4.2 Principal component analysis (PCA) of DNA methylation profiles based on all CpG sites after filtering out low coverage sites (See Methods). HL and KL individuals are separated along the PC1, and the parental and F2 individuals are separated along PC2. PCA axes explain 12.7% (PC1) and 5.1% (PC2) of the total variation. BI_P: marine parental fish. HL_P: freshwater parental fish from HL. KL_P: freshwater parental fish from KL. HL_F1: hybrid F1 fish from HL lines. KL_F1: hybrid F1 fish from KL lines. HL_F2: hybrid F2 fish from HL lines. KL_F2: hybrid F2 fish from KL lines.

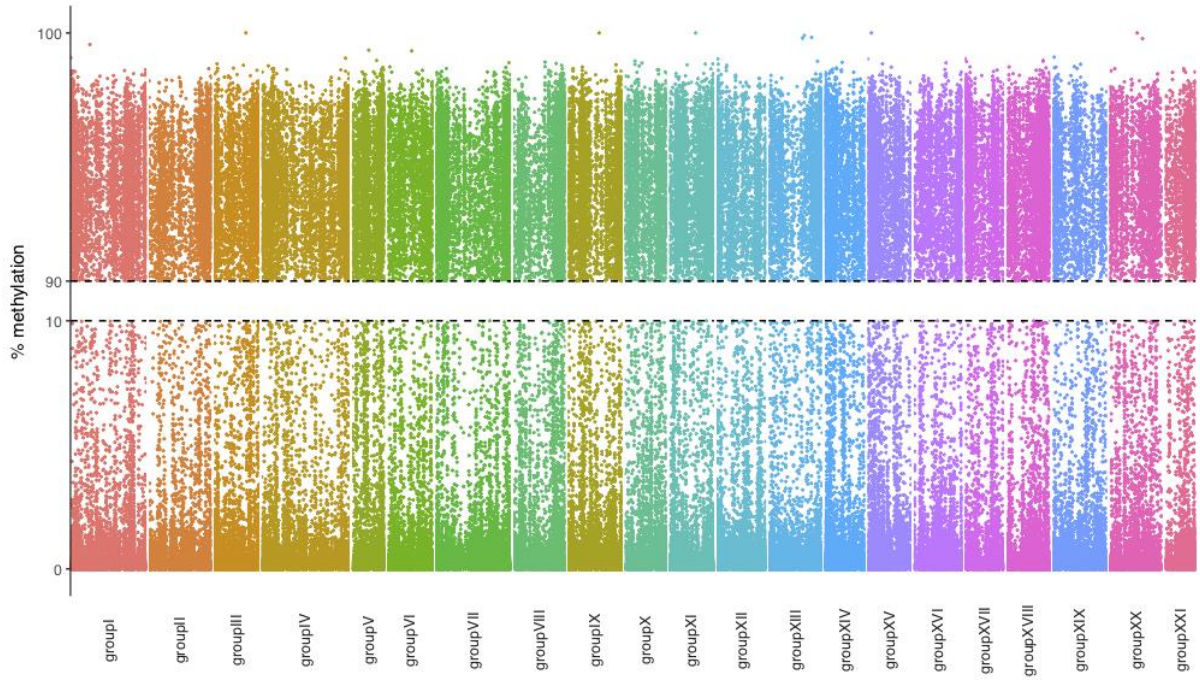


Fig. 4.3 Manhattan plot of the chromosomal positions of constitutively hypo- and hypermethylated CpG sites. HL_F2 is displayed as an example. Each point represents a CpG site. The y-axis presents the average methylation level for that CpG site. Only CpG sites with lower than 10% (constitutively hypomethylation) or higher than 90% (constitutively hypermethylated) methylation level are shown. Points below and above the horizontal dashed lines are constitutively hypo- and hypermethylated loci, respectively.

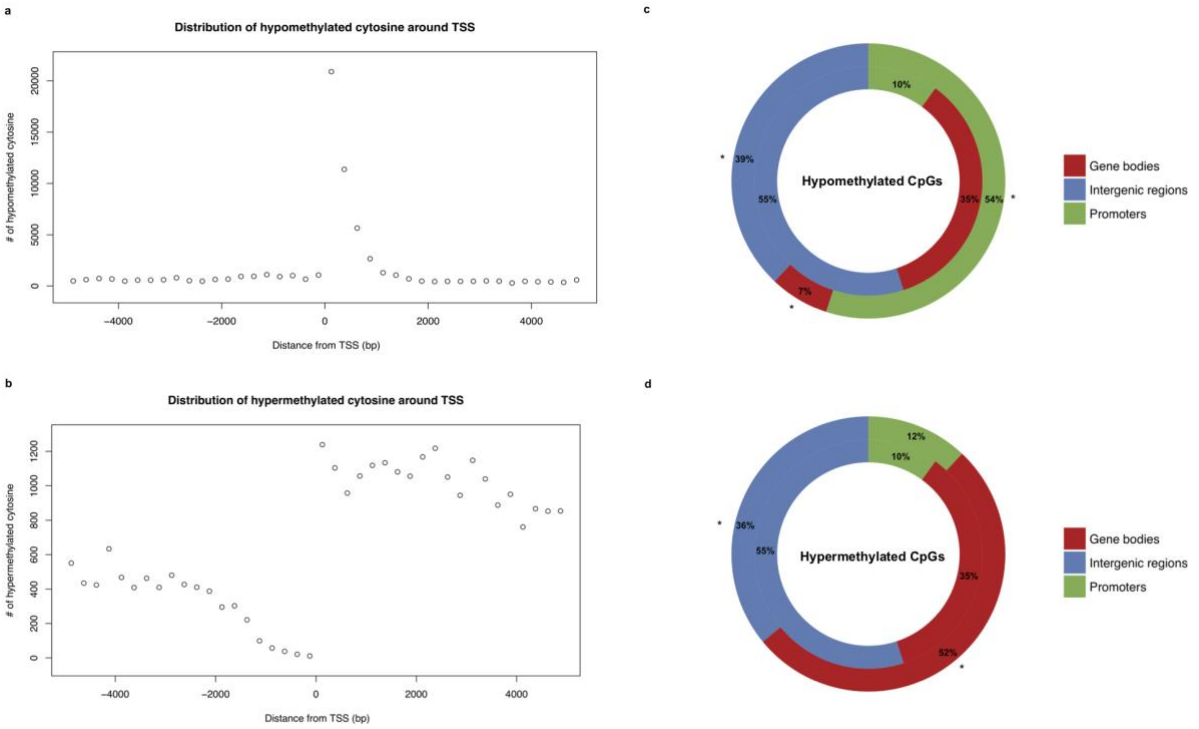


Fig. 4.4 (a-b) The position of (a) constitutively hypo- or (b) hypermethylated CpG sites to the annotated transcription start sites (TSSs) of the *G. aculeatus* genome. (c-d) The proportion of genomic features (promoters, gene bodies, or intergenic regions) in (c) constitutively hypo- or (d) hypermethylated CpG sites compared with the rest of the genome. The outer ring describes the locations of CpG sites, the inner ring describes the genome-wide features. Asterisks denote significant differences between the features of the CpG sites versus the rest of the genome using a *G*-test at $P < 0.01$.

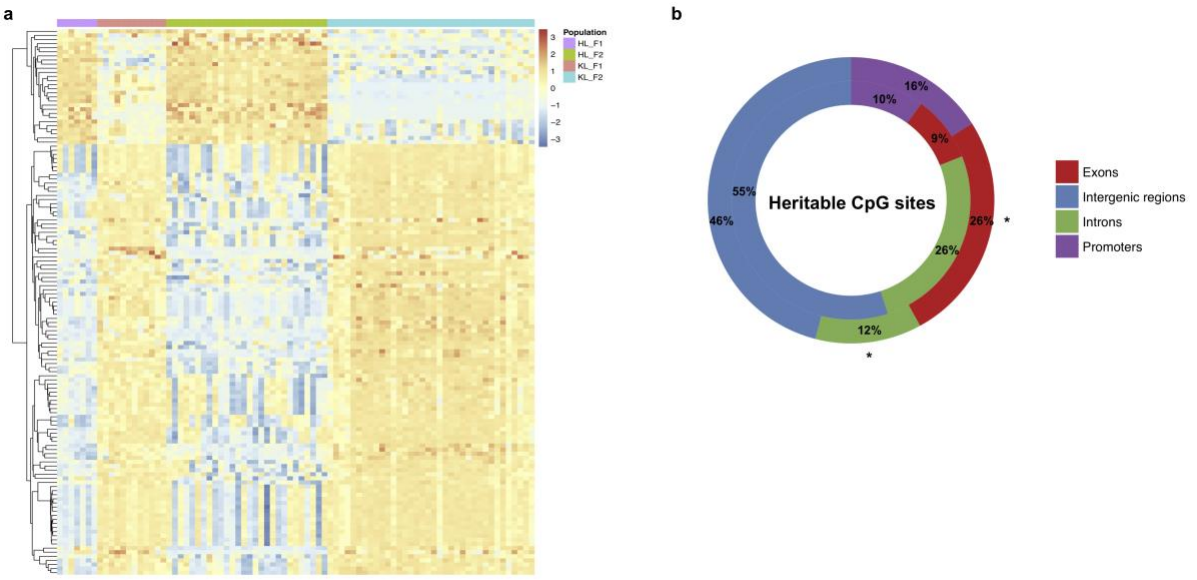


Fig. 4.5 (a) Heatmap of methylation levels of shared DMCs between F1 and F2 generations. Each column represents a colour-coded individual: purple for F1 fish from HL lines (HL_F1), red for F1 fish from KL lines (KL_F1), green for F2 fish from HL lines (HL_F2), and blue for F2 fish from KL lines (KL_F2). Each row represents one of the shared DMCs. DMCs are clustered based on the similarities of the methylation patterns between individuals. The darker the red, the more methylated that individual is for that DMC. The darker the blue, the less methylated that individual is for that DMC. (b) The proportion of genomic features (promoters, exons, introns, or intergenic regions) in DMCs compared with the rest of the genome. The outer ring describes the locations of DMCs, the inner ring describes the genome-wide features. Asterisks denote significant differences between the features of the DMCs versus the rest of the genome using a *G*-test at $P < 0.001$.

References

Akalin A, Franke V, Vlahoviček K, Mason CE, Schübeler D. 2015. genomation: a toolkit to summarize, annotate and visualize genomic intervals. *Bioinformatics* 31:1127-1129.

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

Akimenko MA, Johnson SL, Westerfield M, Ekker M. 1995. Differential induction of four msx homeobox genes during fin development and regeneration in zebrafish. *Development* 121:347.

Alexa A, Rahnenführer J, Lengauer T. 2006. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* 22:1600-1607.

Artemov AV, Mugue NS, Rastorguev SM, Zhenilo S, Mazur AM, Tsygankova SV, Boulygina ES, Kaplun D, Nedoluzhko AV, Medvedeva YA, et al. 2017. Genome-wide DNA methylation profiling reveals epigenetic adaptation of stickleback to marine and freshwater conditions. *Mol Biol Evol.* 34:2203-2213.

Baerwald MR. 2013. Temporal expression patterns of rainbow trout immune-related genes in response to *Myxobolus cerebralis* exposure. *Fish Shellfish Immunol.* 35:965-971.

Baerwald MR, Meek MH, Stephens MR, Nagarajan RP, Goodbla AM, Tomalty KM, Thorgaard GH, May B, Nichols KM. 2015. Migration-related phenotypic divergence is associated with epigenetic modifications in rainbow trout. *Mol Ecol.* 25:1785-1800.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3:e3376.

Bell MA, Foster SA. 1994. The evolutionary biology of the threespine stickleback. New York: Oxford University Press.

Boyle P, Clement K, Gu H, Smith ZD, Ziller M, Fostel JL, Holmes L, Meldrim J, Kelley F, Gnirke A, et al. 2012. Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. *Genome Biol.* 13:R92.

Chatterjee A, Ozaki Y, Stockwell PA, Horsfield JA, Morison IM, Nakagawa S. 2013. Mapping the zebrafish brain methylome using reduced representation bisulfite sequencing. *Epigenetics* 8:979-989.

Daxinger L, Whitelaw E. 2012. Understanding transgenerational epigenetic inheritance via the gametes in mammals. *Nat Rev Genet.* 13:153-162.

Feil R, Fraga MF. 2012. Epigenetics and the environment: emerging patterns and implications. *Nat Rev Genet.* 13:97-109.
Feis ME, John U, Lokmer A, Luttikhuizen PC, Wegner KM. 2018. Dual transcriptomics reveals co-evolutionary mechanisms of intestinal parasite infections in blue mussels *Mytilus edulis*. *Mol Ecol.* 27:1505-1519.
Feng S, Cokus SJ, Zhang X, Chen P-Y, Bostick M, Goll MG, Hetzel J, Jain J, Strauss SH, Halpern ME, et al. 2010. Conservation and divergence of methylation patterning in plants and animals. *Proc Natl Acad Sci U S A.* 107:8689-8694.
Furey TS. 2012. ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions. *Nat Rev Genet.* 13:840-852.
Gibbons TC, Metzger DCH, Healy TM, Schulte PM. 2017. Gene expression plasticity in response to salinity acclimation in threespine stickleback ecotypes from different salinity habitats. *Mol Ecol.* 26:2711-2725.
Gu H, Smith ZD, Bock C, Boyle P, Gnirke A, Meissner A. 2011. Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat Protoc.* 6:468-481.
Gugger PF, Fitz-Gibbon S, Pellegrini M, Sork VL. 2016. Species-wide patterns of DNA methylation variation in Quercus lobata and its association with climate gradients. *Mol Ecol.* 25:1665-1680.
Haase D, Rieger JK, Witten A, Stoll M, Bornberg-Bauer E, Kalbe M, Reusch TB. 2014. Specific gene expression responses to parasite genotypes reveal redundancy of innate immunity in vertebrates. *PLoS ONE* 9:e108001.
Heard E, Martienssen RA. 2014. Transgenerational epigenetic inheritance: myths and mechanisms. *Cell* 157:95-109.
Hellman A, Chess A. 2007. Gene body-specific methylation on the active X chromosome. *Science* 315:1141.
Hibbeler S, Scharsack JP, Becker S. 2008. Housekeeping genes for quantitative expression studies in the three-spined stickleback Gasterosteus aculeatus. *BMC Mol Biol.* 9:18.
Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet.* 6:e1000862.
Hu J, Barrett RDH. 2017. Epigenetics in natural animal populations. *J Evol Biol.* 30:1612-1632.

Huang Y, Chain FJ, Panchal M, Eizaguirre C, Kalbe M, Lenz TL, Samonte IE, Stoll M, Bornberg-Bauer E, Reusch TB, et al. 2016. Transcriptome profiling of immune tissues reveals habitat-specific gene expression between lake and river sticklebacks. *Mol Ecol.* 25:943-958.

Jablonka E, Raz G. 2009. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *Q Rev Biol.* 84:131-176.

Jesus TF, Grosso AR, Almeida-Val VM, Coelho MM. 2016. Transcriptome profiling of two Iberian freshwater fish exposed to thermal stress. *J Therm Biol.* 55:54-61.

Jiang L, Zhang J, Wang J-J, Wang L, Zhang L, Li G, Yang X, Ma X, Sun X, Cai J, et al. 2013. Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. *Cell* 153:773-784.

Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, Johnson J, Swofford R, Pirun M, Zody MC, White S, et al. 2012. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484:55-61.

Jones PA. 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet.* 13:484-492.

Katogi R, Nakatani Y, Shin-i T, Kohara Y, Inohaya K, Kudo A. 2004. Large-scale analysis of the genes involved in fin regeneration and blastema formation in the medaka, *Oryzias latipes*. *Mech Dev.* 121:861-872.

Klironomos FD, Berg J, Collins S. 2013. How epigenetic mutations can affect genetic evolution: model and mechanism. *Bioessays* 35:571-578.

Kratochwil CF, Meyer A. 2015. Mapping active promoters by ChIP-seq profiling of H3K4me3 in cichlid fish - a first step to uncover cis-regulatory elements in ecological model teleosts. *Mol Ecol Resour* 15:761-771.

Krueger F, Andrews S. 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 27:1571-1572.

Lam LL, Emberly E, Fraser HB, Neumann SM, Chen E, Miller GE, Kobor MS. 2012. Factors underlying variable DNA methylation in a human community cohort. *Proc Natl Acad Sci U S A.* 109:17253-17260.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Meth.* 9:357-359.

Laurent L, Wong E, Li G, Huynh T, Tsirigos A, Ong CT, Low HM, Kin Sung KW, Rigoutsos I, Loring J, et al. 2010. Dynamic changes in the human methylome during differentiation. *Genome Res.* 20:320-331.

Lawrence M, Huber W, Pages H, Aboyoun P, Carlson M, Gentleman R, Morgan MT, Carey VJ. 2013. Software for computing and annotating genomic ranges. *PLoS Comput Biol.* 9:e1003118.
Le Luyer J, Laporte M, Beacham TD, Kaukinen KH, Withler RE, Leong JS, Rondeau EB, Koop BF, Bernatchez L. 2017. Parallel epigenetic modifications induced by hatchery rearing in a Pacific salmon. *Proc Natl Acad Sci U S A.* 114:12964-12969.
Lea AJ, Altmann J, Alberts SC, Tung J. 2016. Resource base influences genome-wide DNA methylation levels in wild baboons (*Papio cynocephalus*). *Mol Ecol.* 25:1681-1696.
Lim JP, Brunet A. 2013. Bridging the transgenerational gap with epigenetic memory. *Trends Genet.* 29:176-186.
Lister R, Pelizzola M, Dowen R, Hawkins R, Hon G, Tonti-Filippini J, Nery J, Lee L, Ye Z, Ngo Q-M, et al. 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462:315 - 322.
Liu Y, Siegmund KD, Laird PW, Berman BP. 2012. Bis-SNP: combined DNA methylation and SNP calling for Bisulfite-seq data. *Genome Biol.* 13:R61.
López-Maury L, Marguerat S, Bahler J. 2008. Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nat Rev Genet.* 9:583-593.
Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, Johnson BE, Hong C, Nielsen C, Zhao Y, et al. 2010. Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* 466:253-257.
Meissner A, Mikkelsen T, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein B, Nusbaum C, Jaffe D, et al. 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454:766 - 770.
Metzger DCH, Schulte PM. 2018. The DNA methylation landscape of stickleback reveals patterns of sex chromosome evolution and effects of environmental salinity. *Genome Biol Evol.* 10:775-785.
Metzger DCH, Schulte PM. 2017. Persistent and plastic effects of temperature on DNA methylation across the genome of threespine stickleback (*Gasterosteus aculeatus*). *P R Soc B* 284:20171667.
Mogi M, Yokoi H, Suzuki T. 2017. Analyses of the cellular clock gene expression in peripheral tissue, caudal fin, in the Japanese flounder, *Paralichthys olivaceus*. *Gen Comp Endocrinol.* 248:97-105.

Nätt D, Rubin C-J, Wright D, Johnsson M, Beltéky J, Andersson L, Jensen P. 2012. Heritable genome-wide variation of gene expression and promoter methylation between wild and domesticated chickens. *BMC Genomics* 13:59.
O'Dea RE, Noble DWA, Johnson SL, Hesselson D, Nakagawa S. 2016. The role of non-genetic inheritance in evolutionary rescue: epigenetic buffering, heritable bet hedging and epigenetic traps. *Environmental epigenetics* 2.
Offen N, Meyer A, Begemann G. 2009. Identification of novel genes involved in the development of the sword and gonopodium in swordtail fish. *Dev Dyn.* 238:1674-1687.
Ohtsubo K, Marth JD. 2006. Glycosylation in cellular mechanisms of health and disease. *Cell* 126:855-867.
Orozco LD, Morselli M, Rubbi L, Guo W, Go J, Shi H, Lopez D, Furlotte NA, Bennett BJ, Farber CR, et al. 2015. Epigenome-wide association of liver methylation patterns and complex metabolic traits in mice. *Cell Metab.* 21:905-917.
Park PJ. 2009. ChIP-Seq: advantages and challenges of a maturing technology. *Nat Rev Genet.* 10:669-680.
Peichel CL, Marques DA. 2017. The genetic and molecular architecture of phenotypic diversity in sticklebacks. *Philos Trans R Soc Lond B Biol Sci.* 372.
Peichel CL, Ross JA, Matson CK, Dickson M, Grimwood J, Schmutz J, Myers RM, Mori S, Schluter D, Kingsley DM. 2004. The master sex-determination locus in threespine sticklebacks is on a nascent Y chromosome. *Curr Biol* 14:1416-1424.
Potok ME, Nix DA, Parnell TJ, Cairns BR. 2013. Reprogramming the maternal zebrafish genome after fertilization to match the paternal methylation pattern. *Cell* 153:759-772.
Radtke F, Fasnacht N, MacDonald HR. 2010. Notch signaling in the immune system. *Immunity* 32:14-27.
Roberts SB, Gavery MR. 2012. Is there a relationship between DNA methylation and phenotypic plasticity in invertebrates? *Front Physiol.* 2:116.
Rohland N, Reich D. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Res.* 22:939-946.
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.
Shea N, Pen I, Uller T. 2011. Three epigenetic information channels and their different roles in evolution. *J Evol Biol.* 24:1178-1187.

Smith G, Smith C, Kenny JG, Chaudhuri RR, Ritchie MG. 2015. Genome-wide DNA methylation patterns in wild samples of two morphotypes of threespine stickleback (*Gasterosteus aculeatus*). *Mol Biol Evol.* 32:888-895.
Smith ZD, Meissner A. 2013. DNA methylation: roles in mammalian development. *Nat Rev Genet.* 14:204-220.
Storey JD, Tibshirani R. 2003. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A.* 100:9440-9445.
Stricker SH, Koferle A, Beck S. 2017. From profiles to function in epigenomics. *Nat Rev Genet.* 18:51-66.
Suzuki MM, Bird A. 2008. DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet.* 9:465-476.
Taudt A, Colome-Tatche M, Johannes F. 2016. Genetic sources of population epigenomic variation. *Nat Rev Genet.* 17:319-332.
Tu S, Johnson Stephen L. 2011. Fate restriction in the growing and regenerating zebrafish fin. *Dev Cell* 20:725-732.
van Eeden FJ, Granato M, Schach U, Brand M, Furutani-Seiki M, Haffter P, Hammerschmidt M, Heisenberg CP, Jiang YJ, Kane DA, et al. 1996. Genetic analysis of fin formation in the zebrafish, *Danio rerio*. *Development* 123:255.
Verhoeven KJF, vonHoldt BM, Sork VL. 2016. Epigenetics in ecology and evolution: what we know and what we need to know. *Mol Ecol.* 25:1631-1638.
Walker JA. 2008. Ecological morphology of lacustrine threespine stickleback *Gasterosteus aculeatus* L. (Gasterosteidae) body shape. *Biol J Linn Soc.* 61:3-50.
Weyrich A, Lenz D, Jeschek M, Chung TH, Rubensam K, Goritz F, Jewgenow K, Fickel J. 2016. Paternal intergenerational epigenetic response to heat exposure in male Wild guinea pigs. *Mol Ecol.* 25:1729-1740.
Wittkopp PJ, Kalay G. 2011. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat Rev Genet.* 13:59.
Wolf SF, Jolly DJ, Lunnen KD, Friedmann T, Migeon BR. 1984. Methylation of the hypoxanthine phosphoribosyltransferase locus on the human X chromosome: implications for X-chromosome inactivation. *Proc Natl Acad Sci U S A.* 81:2806-2810.
Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW. 2010. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet.* 42:565.

Yoshida S, Date Y, Akama M, Kikuchi J. 2014. Comparative metabolomic and ionomic approach for abundant fishes in estuarine environments of Japan. *Sci Rep.* 4:7005.
Zhu LJ. 2013. Integrative analysis of ChIP-Chip and ChIP-Seq dataset. In: Lee T-L, Shui Luk AC, editors. Tiling Arrays: Methods and Protocols. Totowa, NJ: Humana Press. p. 105-124.
Zhu LJ, Gazin C, Lawson ND, Pagès H, Lin SM, Lapointe DS, Green MR. 2010. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics* 11:237.

Table S4.1 Over-represented GO terms of genes overlapping with constitutively hypo- or hypermethylated CpG sites

GO ID	Term	*P*-value	FDR	Hypo/Hyper	Category†
GO:0000188	Inactivation of MAPK activity	2.70E-04	1.38E-02	Hypo	BP
GO:0048596	Embryonic camera-type eye morphogenesis	2.70E-04	1.38E-02	Hypo	BP
GO:0021520	Spinal cord motor neuron cell fate specification	5.30E-04	2.54E-02	Hypo	BP
GO:0001731	Formation of translation preinitiation complex	8.90E-04	3.87E-02	Hypo	BP
GO:0006357	Regulation of transcription from RNA polymerase II	6.60E-05	4.47E-03	Hypo	BP
GO:0006355	Regulation of transcription, DNA-templated	8.30E-07	9.47E-05	Hypo	BP
GO:0006412	Translation	7.50E-07	8.76E-05	Hypo	BP
GO:0010506	Regulation of autophagy	8.90E-04	3.87E-02	Hypo	BP
GO:0060038	Cardiac muscle cell proliferation	5.30E-04	2.54E-02	Hypo	BP
GO:0061371	Determination of heart left/right asymmetry	3.90E-04	1.96E-02	Hypo	BP
GO:0007275	Multicellular organism development	4.40E-16	1.24E-12	Hypo	BP
GO:0043009	Chordate embryonic development	1.10E-04	6.66E-03	Hypo	BP
GO:0001654	Eye development	1.90E-04	1.07E-02	Hypo	BP
GO:0055114	Oxidation-reduction process	6.70E-04	3.16E-02	Hypo	BP
GO:0007399	Nervous system development	3.00E-07	3.83E-05	Hypo	BP
GO:0071901	Negative regulation of protein serine/threonine kinase activity	1.30E-05	1.04E-03	Hypo	BP
GO:0031076	Embryonic camera-type eye development	4.20E-05	3.00E-03	Hypo	BP
GO:0008152	Metabolic process	9.10E-09	2.72E-06	Hypo	BP
GO:0007420	Brain development	1.60E-04	8.94E-03	Hypo	BP
GO:0006351	Transcription, DNA-templated	2.30E-07	3.05E-05	Hypo	BP
GO:0006413	Translational initiation	9.90E-04	4.16E-02	Hypo	BP
GO:0030097	Hemopoiesis	3.40E-05	2.50E-03	Hypo	BP
GO:0007507	Heart development	2.40E-06	2.39E-04	Hypo	BP
GO:0009887	Animal organ morphogenesis	3.30E-05	2.43E-03	Hypo	BP
GO:0030182	Neuron differentiation	7.20E-06	6.32E-04	Hypo	BP
GO:0001708	Cell fate specification	8.30E-05	5.41E-03	Hypo	BP
GO:0005634	Nucleus	2.00E-07	2.00E-05	Hypo	CC
GO:0005840	Ribosome	6.00E-05	3.20E-03	Hypo	CC
GO:0005852	Eukaryotic translation	6.00E-04	2.26E-02	Hypo	CC

	initiation factor 3 complex				
GO:0016282	Eukaryotic 43S preinitiation complex	6.00E-04	2.26E-02	Hypo	CC
GO:0033290	Eukaryotic 48S preinitiation complex	6.00E-04	2.26E-02	Hypo	CC
GO:0005622	Intracellular	3.70E-15	2.75E-12	Hypo	CC
GO:0005739	Mitochondrion	7.90E-05	3.95E-03	Hypo	CC
GO:0005623	Cell	2.10E-07	2.00E-05	Hypo	CC
GO:0043231	Intracellular membrane-bounded organelle	1.20E-10	2.32E-08	Hypo	CC
GO:0005737	Cytoplasm	6.30E-06	3.63E-04	Hypo	CC
GO:0030529	Intracellular ribonucleoprotein complex	5.10E-06	3.22E-04	Hypo	CC
GO:0043565	Sequence-specific DNA binding	3.30E-05	2.20E-02	Hypo	MF
GO:0008536	Ran GTPase binding	9.70E-05	3.53E-02	Hypo	MF
GO:0003735	Structural constituent of ribosome	1.10E-04	3.53E-02	Hypo	MF
GO:0017017	MAP kinase tyrosine/serine/threonine phosphatase activity	2.10E-04	4.15E-02	Hypo	MF
GO:0019843	rRNA binding	7.50E-05	3.53E-02	Hypo	MF
GO:0003677	DNA binding	9.50E-06	9.49E-03	Hypo	MF
GO:0043547	Positive regulation of GTPase activity	4.40E-05	7.11E-03	Hyper	BP
GO:0006486	Protein glycosylation	2.60E-06	7.10E-04	Hyper	BP
GO:0018108	Peptidyl-tyrosine phosphorylation	2.10E-04	2.47E-02	Hyper	BP
GO:0048747	Muscle fiber development	1.30E-04	1.78E-02	Hyper	BP
GO:0007507	Heart development	2.80E-05	5.23E-03	Hyper	BP
GO:0043009	Chordate embryonic development	3.10E-04	3.31E-02	Hyper	BP
GO:0048840	Otolith development	4.40E-04	4.29E-02	Hyper	BP
GO:0008593	Regulation of Notch signaling pathway	3.60E-04	3.74E-02	Hyper	BP
GO:0007219	Notch signaling pathway	6.00E-05	9.49E-03	Hyper	BP
GO:0050793	Regulation of developmental process	2.40E-04	2.69E-02	Hyper	BP
GO:0043087	Regulation of GTPase activity	1.50E-05	3.09E-03	Hyper	BP
GO:0005622	Intracellular	1.30E-06	3.20E-04	Hyper	CC
GO:0005623	Cell	2.70E-07	2.00E-04	Hyper	CC
GO:0005654	Nucleoplasm	4.90E-04	4.84E-02	Hyper	CC
GO:0005524	ATP binding	4.20E-09	2.83E-06	Hyper	MF
GO:0005096	GTPase activator activity	6.10E-05	7.67E-03	Hyper	MF
GO:0046872	Metal ion binding	9.60E-07	3.22E-04	Hyper	MF
GO:0004222	Metalloendopeptidase activity	2.70E-04	2.13E-02	Hyper	MF
GO:0018024	Histone-lysine N-methyltransferase activity	3.40E-05	5.61E-03	Hyper	MF

GO:0005085	Guanyl-nucleotide exchange factor activity	4.60E-04	3.18E-02	Hyper	MF
GO:0000166	Nucleotide binding	3.00E-04	2.23E-02	Hyper	MF
GO:0004672	Protein kinase activity	2.10E-04	1.93E-02	Hyper	MF
GO:0030695	GTPase regulator activity	6.80E-05	8.05E-03	Hyper	MF
GO:0004713	Protein tyrosine kinase activity	6.70E-04	4.01E-02	Hyper	MF
GO:0016887	ATPase activity	2.50E-04	2.05E-02	Hyper	MF

† BP: Biological Process; CC: Cellular Component; MF: Molecular Function

Table S4.2 List of genes mapped by heritable methylated CpG sites

Symbol	Description
adam11	ADAM metallopeptidase domain 11
arel1	Apoptosis resistant E3 ubiquitin protein ligase 1
arhgef26	Rho guanine nucleotide exchange factor 26
dhrs13	Dehydrogenase/reductase 13
dst	Dystonin
galnt10	Polypeptide N-acetylgalactosaminyltransferase 10
hist1h2bm	Histone cluster 1 H2B
mmp16	Matrix metallopeptidase 16
pok3r5	Phosphoinositide-3-kinase regulatory subunit 5
ryr3	Ryanodine receptor 3
spag9	Sperm associated antigen
aatf	Apoptosis antagonizing transcription factor
abcg2d	ATP-binding cassette, sub-family G (WHITE), member 2d
adamts14	ADAM metallopeptidase with thrombospondin type 1 motif 14
atp2b2	ATPase plasma membrane Ca2+ transporting 2
clip2	CAP-GLY domain containing linker protein 2
clvs2	Clavesin 2
dpp6a	Dipeptidyl-peptidase 6a
gyg1a	Glycogenin 1a
htr1d	5-hydroxytryptamine (serotonin) receptor 1D, G protein-coupled
itih5	Inter-alpha-trypsin inhibitor heavy chain family, member 5
lats2	Iarge tumor suppressor kinase 2
lin9	Lin-9 DREAM MuvB core complex component
loxl5b	Lysyl oxidase-like 5b
pcdh7b	Protocadherin 7b
pdxkb	Pyridoxal (pyridoxine, vitamin B6) kinase b
psmd1	Proteasome 26S subunit, non-ATPase 1
secisbp2	SECIS binding protein 2
sema6e	Sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6E
sirt7	Sirtuin 7
ttn.2	Titin, tandem duplicate 2
znf385b	Zinc finger protein 385B

GENERAL CONCLUSION

Overview

One of the recurring questions in ecology and evolution is: what is the mechanism underlying phenotypic variation? The prevailing theory is that phenotypic variation originates from genetic and environmental variation, and/or their interactions. However, in the case of insufficient genetic variation, environmental variation can be the main driver for the phenotypic variation observed in some animal populations, whereas epigenetics can play a role in mediating such phenotypic response to the environment. Furthermore, phenotypic change can occur in sufficiently short time scales that genetic modification may not be the sole mechanism underlying these interactions. An increasing number of recent studies have suggested that, in addition to genetic variation, epigenetic variation can also contribute to phenotypic variation, as well as some evolutionary processes. Here, I have characterised epigenetic variation in three distinct empirical systems, and investigated its role in phenotypic variation and development. My overarching goal has been to expand our knowledge of the mechanisms underlying phenotypic variation and highlight the importance of epigenetic variation for evolution.

General summary

I first reviewed the literature on ecological epigenetics in natural animal populations. While previous studies have characterised epigenetic responses to environmental change in lab-reared animals, such responses under laboratory conditions may not reflect outcomes observed under natural settings, due to the higher genetic heterogeneity in natural populations and greater levels of environmental complexity in nature. I provided an overview of recent empirical studies in natural animal populations, and found that epigenetic variation can be shaped by the environment, resulting in phenotypic variation. In addition, epigenetic variation may contribute to the adaptive differentiation between populations residing in different environments. I also reviewed theoretical studies exploring the role of epigenetic variation in evolutionary processes. Depending on the autonomy of epigenetic variation from genetic variation, and the frequency of environmental change, epigenetic variation can facilitate faster exploration of fitness landscapes, and contribute to the evolutionary process. My review, coupled with existing knowledge about epigenetics in ecology and evolution, provides a more complete picture of how epigenetic variation acts in natural populations.

To further our understanding of the role of epigenetic variation, especially the effects of DNA methylation on phenotypic variation and development in natural animal populations, I then profiled genome-wide methylation patterns using reduced representation bisulphite sequencing in three distinct empirical systems. In my second chapter, I compared DNA methylation variation in guppies at different infection phases, and found that hypermethylation was a general response to parasite infection. I also found unique responses at distinct infection phases, with several differentially methylated regions (DMRs) overlapping with immune genes. Results from this chapter add to previous work on epigenetic responses to endoparasitic infection by providing the first study of an epigenetic signature of infection by ectoparasites, and demonstrate the relationship between epigenetic variation and immune response in distinct phases of infection.

To next investigate epigenetic responses under natural conditions, my third chapter focused on methylation responses in *Anolis* lizards transplanted from a source population to either high- or low-quality islands. Surprisingly, I again found hypermethylation was a general response to habitat change, with the magnitude of increased methylation negatively associated with habitat quality. Identified DMRs mapped to genes relevant to environmental change, e.g., neuronal modification, skeletal and muscle growth, and immune response, suggesting a relationship between environmental change and epigenetically regulated phenotypes at the early stage of colonization.

My previous two chapters focused on the methylation variation that can be altered by environmental factors; however, most genomic methylation patterns are somatically static across tissues and throughout life time in vertebrates. In addition, the amount of heritable DNA methylation has typically been assumed to be minimal due to epigenetic resetting during embryogenesis. Yet, few studies have actually explored heritable methylation in natural animal populations. To address this gap, my last chapter characterized 1) general methylation patterns across generations, 2) constitutively methylated regions and their functions, and 3) the amount of heritable methylation and its genetic basis in threespine stickleback. I found that general methylation patterns were associated with genetic relatedness. In addition, I found that constitutive methylation was common, and some of the implicated regions overlapped with genes related to structural maintenance, and fin development and regeneration. Furthermore, I discovered highly consistent methylation patterns across generations, suggesting that the epigenetic resetting is not fully complete in sticklebacks, and that the loci associated with these patterns may have evolutionary importance. This chapter provides the first study of constitutive and heritable methylation in

natural populations of fish, demonstrates the importance of constitutive methylation as an epigenetic regulatory mechanism, and improves our understanding of the heritable basis of population epigenomic variation.

Implications

My dissertation helps to increase our knowledge of the relationship between epigenetics, phenotypic plasticity, and evolution. While the study of epigenetics in ecology and evolution is still at its infancy, our understanding of epigenetics is increasing due to mounting evidence suggesting its relevance for altered gene expression and phenotypic responses to changing environments. With increasing anthropogenic disturbances in ecosystems leading to increasing environmental fluctuations, labile variation such as epigenetic variation could allow for faster response to such changes. Yet, at the same time, epigenetics can contribute to the maintenance of the normal development of tissues. Thus, my dissertation underscores the importance of taking epigenetic variation into consideration when analyzing animal responses to environmental change.