Spring 2021 – Epigenetics and Systems Biology
Discussion Session (Epigenetics and Evolutionary Biology)
Michael K. Skinner – Biol 476/576
Week 15 (April 29)

Epigenetics and Evolutionary Biology

Primary Papers

1. Luo, et al. (2020) Cell Reports 33:108306. (PMID: 33113358)
2. Heckwolf, et al. (2020) Science Advances 6:eaaz1138. (PMID: 32219167)
3. McNew, et al. (2017) BMC Evolution 17(1):183. (PMID: 28835203)

Discussion

Student 41 – Ref #1 above

- What was the experimental design and model system?
- What epigenetic process and gene network effects were observed?
- Does this provide evidence for environmental induction of epigenetic alterations in a gene network for evolutionary adaptation?

Student 42 – Ref #2 above

- What is the model system, phenotypic change, and environmental factor?
- What epigenetic change was observed?
- How did the environment, epigenetics and genetics integrate?

Student 43 – Ref #3 above

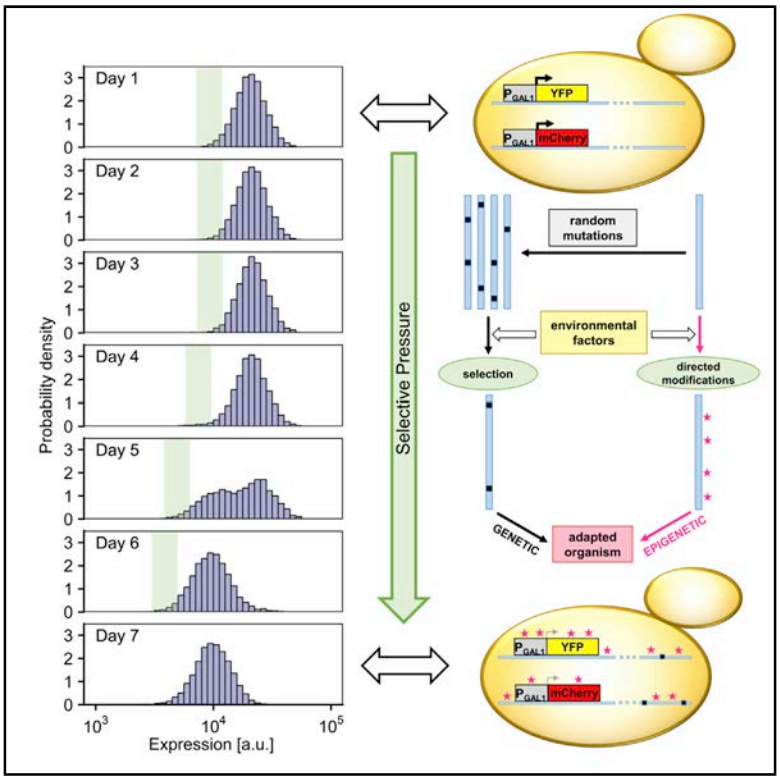
- What was the experimental design and approach?
- What molecular alterations were observed in what cell types?
- What molecular mechanism can promote rapid evolutionary events?

Cell Reports

Article

Epigenetic Mechanisms Contribute to Evolutionary Adaptation of Gene Network Activity under Environmental Selection

Graphical Abstract



Authors

Xinyue Luo, Ruijie Song, David F. Moreno, Hong-Yeoul Ryu, Mark Hochstrasser, Murat Acar

Correspondence

murat.acar@yale.edu

In Brief

Luo et al. demonstrate how epigenetic mechanisms contribute to the evolution of gene network activity. Subjecting yeast cells to repeated environmental selection based on the activity of the galactose network, they observe sustained changes in reporter expression level. They characterize the epigenetic and genetic factors contributing to the observed phenotypes.

Highlights

- Gated sorting and regrowth of low-expression cells causes sustained expression decrease
- Chromatin environment affects expression changes in response to selective pressure
- Genetic causes alone cannot explain the expression reduction under selection
- Inheritance of epigenetic factors contributes to the observed phenotypic change

Luo et al., 2020, Cell Reports *33*, 108306
October 27, 2020
https://doi.org/10.1016/j.celrep.2020.108306

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Epigenetic Mechanisms Contribute to Evolutionary Adaptation of Gene Network Activity under Environmental Selection

Xinyue Luo,[1,2,6] Ruijie Song,[2,3,6] David F. Moreno,[1,2] Hong-Yeoul Ryu,[4,7] Mark Hochstrasser,[1,4] and Murat Acar[1,2,3,5,8,*]
[1]Department of Molecular Cellular and Developmental Biology, Yale University, 219 Prospect Street, New Haven, CT 06511, USA
[2]Systems Biology Institute, Yale University, 850 West Campus Drive, West Haven, CT 06516, USA
[3]Interdepartmental Program in Computational Biology and Bioinformatics, Yale University, 300 George Street, Suite 501, New Haven, CT 06511, USA
[4]Department of Molecular Biophysics and Biochemistry, Yale University, 266 Whitney Avenue, New Haven, CT 06520, USA
[5]Department of Physics, Yale University, 217 Prospect Street, New Haven, CT 06511, USA
[6]These authors contributed equally
[7]Present address: School of Life Sciences, BK21 Plus KNU Creative BioResearch Group, College of National Sciences, Kyungpook National University, Daegu 41566, Republic of Korea
[8]Lead Contact
*Correspondence: murat.acar@yale.edu
https://doi.org/10.1016/j.celrep.2020.108306

SUMMARY

How evolution can be facilitated by epigenetic mechanisms has received refreshed attention recently. To explore the role epigenetic inheritance plays in evolution, we subject isogenic wild-type yeast cells expressing P_{GAL1}-YFP (yellow fluorescent protein) to selection by daily sorting based on reporter expression. We observe expression-level reductions in multiple replicates sorted for the lowest expression that persist for several days, even after lifting the selection pressure. Reduced expression is due to factors in the galactose (GAL) network rather than global factors. Results using a constitutively active GAL network are in overall agreement with findings with the wild-type network. We find that the local chromatin environment of the reporter has a significant effect on the observed phenotype. Genome sequencing, chromatin immunoprecipitation (ChIP)-qPCR, and sporulation analysis provide further insights into the epigenetic and genetic contributors to the expression changes observed. Our work provides a comprehensive example of the role played by epigenetic mechanisms on gene network evolution.

INTRODUCTION

Since Charles Darwin published *On the Origin of Species* in 1859, the concept of evolution by natural selection has occupied a prominent place in modern biology (Darwin, 1859). Darwin himself, of course, had no knowledge of the molecular details of this process: DNA would not be established as the genetic material for another 85 years (Avery et al., 1944). The neo-Darwinian evolution theory combines modern knowledge of genetics and molecular biology with Darwin's thinking (Olson-Manning et al., 2012), but classic neo-Darwinian evolution theory is focused on genetics as the primary molecular mechanism and has substantial difficulties with the fact that beneficial genetic mutations occur at an extremely low rate (Day and Bonduriansky, 2011; Jablonka and Raz, 2009; Kuzawa and Thayer, 2011; Nei and Nozawa, 2011), to the point where some evolutionary biologists have called for a rethinking of the entire evolutionary theory (Laland et al., 2014).

The concept of inheritance of acquired characteristics is frequently attributed to Lamarck (Skinner, 2015), though perhaps inaccurately (Burkhardt, 2013). Nonetheless, the so-called neo-Lamarckian theory, grounded on epigenetic mechanisms, has received increased attention recently (Burggren, 2014; Day and Bonduriansky, 2011; Skinner et al., 2015). A key postulate of the neo-Lamarckian theory is that environment directly alters phenotype generationally (Figure 1A) (Skinner, 2015). In this context, epigenetic mechanisms can be the mediator for the environment to directly alter phenotypic variation and its subsequent inheritance (Skinner, 2015).

Evolutionary consequences of epigenetic inheritance have been studied in recent years, showing how epigenetic control of gene expression affects adaptation (Bódi et al., 2017; Bonduriansky and Day, 2009; Bonduriansky et al., 2012; Halfmann et al., 2012; Klironomos et al., 2013; Kronholm and Collins, 2016; Stajic et al., 2019). Nongenetic inheritance can be mediated in several ways, such as by the inheritance of epigenetic states, cytoplasmic factors, and nutrients (Bonduriansky and Day, 2009). Nongenetic inheritance and its evolutionary implications have been conceptualized in a general framework, showing that by decoupling phenotypic change from the genotype, nongenetic inheritance could circumvent the limitations of genetic inheritance (Bonduriansky and Day, 2009). Nongenetic and genetic inheritance mechanisms

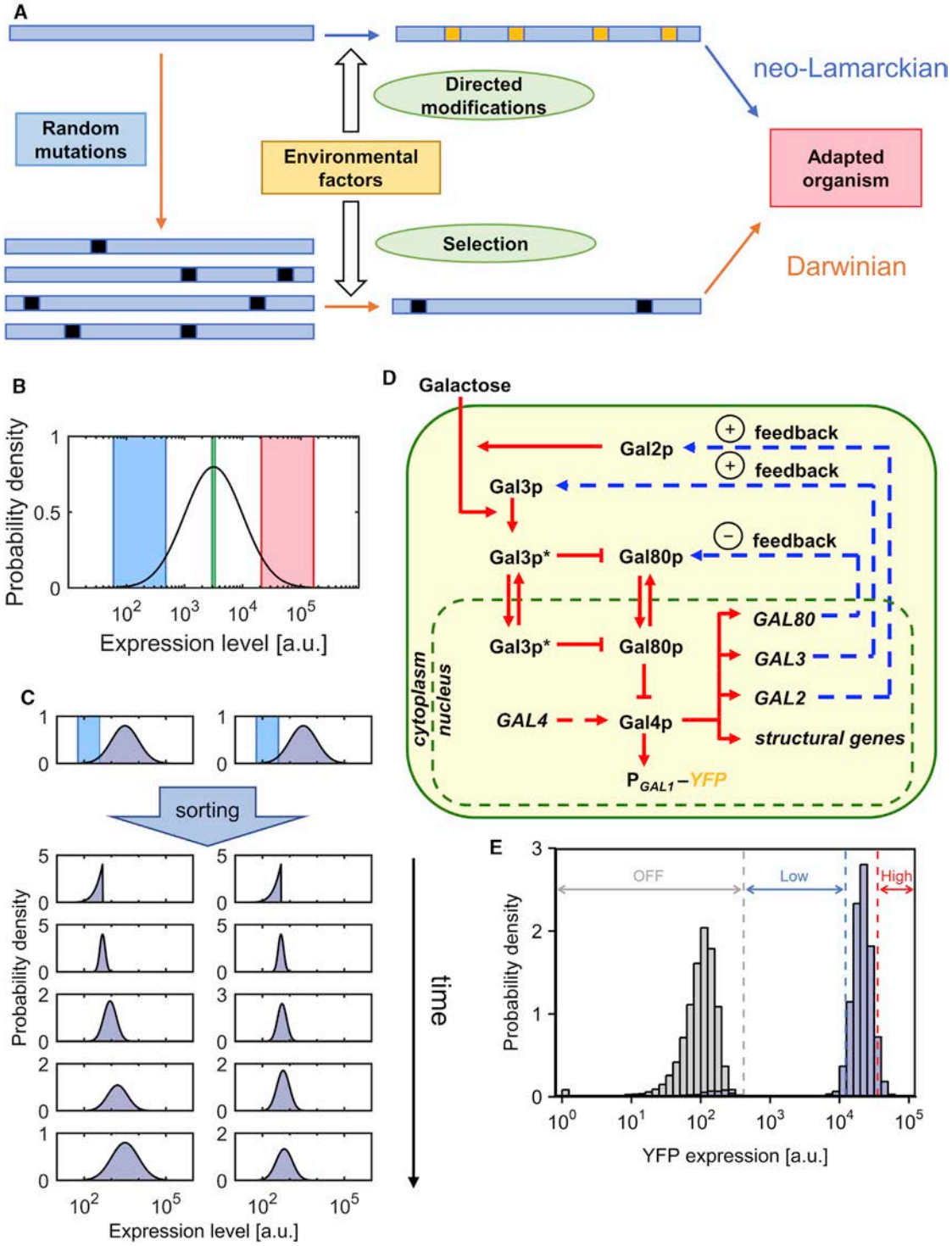


Figure 1. Evolutionary Models, Study Design, and Model Network

(A) Illustration of Darwinian (orange arrows) and neo-Lamarckian (blue arrows) models of evolution. Blue rectangles represent DNA. Black and yellow squares represent genetic mutations and epigenetic modifications, respectively.

(B) Illustration of the three possible sorting gates used: lowest 5% (blue), middle 5% (green), and highest 5% (red).

(C) Illustration of two possible outcomes of the sorting experiment. After being sorted for the lowest-expressing cells, the initial sharp distribution of expression levels will gradually relax, but it may relax either back to the same distribution as the original (left) or to a different distribution with a lower mean (right).

(legend continued on next page)

are not mutually exclusive. For example, theoretical predictions suggested that nongenetic inheritance could increase the rate of both phenotypic and genetic change (Bonduriansky et al., 2012). Also, theoretical and computational work showed how the interplay of heritable epigenetic changes with genetic changes could affect adaptive evolution (Klironomos et al., 2013) and how the effect of epigenetic mutations on adaptive walks depended on their stability and fitness effects relative to genetic mutations (Kronholm and Collins, 2016).

In addition to this theoretical work, experimental studies further focused on the evolutionary consequences of nongenetic heterogeneity and inheritance across generations (Acar et al., 2005, 2008; Bódi et al., 2017; Chatterjee and Acar, 2018; Halfmann et al., 2012; Huang, 2009; Peng et al., 2015; Stajic et al., 2019; Tyedmers et al., 2008; Xue and Acar, 2018a, 2018b). For example, yeast prion proteins can act as epigenetic elements of inheritance (Halfmann et al., 2012), and it has been hypothesized that the yeast prion [*PSI*$^{+}$] provides a mechanism to increase survival in fluctuating environments (Tyedmers et al., 2008); it has also been shown that prions are a common mechanism for phenotypic inheritance in wild strains of *Saccharomyces* (Halfmann et al., 2012). In another experimental study focusing on heterogeneity, it has been shown that phenotypic heterogeneity facilitates adaptive evolution, with the heterogeneity being an evolving trait when populations are under chronic selection pressure (Bódi et al., 2017). As the final example, when tuning low, intermediate, and high levels of heritable silencing of a reporter under selection by insertion within silent subtelomeric yeast chromatin, epigenetic gene silencing has been found to alter the mechanisms and rate of evolutionary adaptation (Stajic et al., 2019).

The concepts of epigenetic inheritance and memory are tightly linked and often used interchangeably to refer to non-DNA-based inheritance (Bonduriansky and Day, 2009). While epigenetic inheritance refers to the passage of certain epigenetic marks to the offspring (Lacal and Ventura, 2018), epigenetic memory is defined as the process of establishing and maintaining a heritable transcriptional state (Acar et al., 2005; Kaufmann et al., 2007; Kundu et al., 2007). Work from the van Oudenaarden group (Acar et al., 2005) described transcriptional memory in the yeast galactose (GAL) network by showing that yeast cells "remember" whether they were previously exposed to high or low concentrations of galactose. Using an engineered GAL network (Acar et al., 2005) where single yeast cells switch between "on" and "off" states of the network, another study (Kaufmann et al., 2007) from the same group measured inheritance of the dynamic gene-expression state and found that several generations after cells have separated, many closely related cell pairs switched with high degrees of synchrony. Providing mechanistic insights into the mediation of epigenetic memory in the GAL network, one study (Kundu et al., 2007) showed that the rate of transcriptional induction of *GAL1* was regulated by the prior expression state; the epigenetic state was inherited by daughter cells, and the SWI/SNF chromatin remodeling enzyme was essential for *GAL1* epigenetic memory. Another study (Brickner et al., 2007) demonstrated that the yeast *GAL1* gene is recruited to the nuclear periphery upon transcriptional activation, and it remains at the periphery for generations after it is repressed, with localization at the periphery serving as a form of memory of recent transcriptional activation. Finally, Tzamarias and colleagues (Zacharioudakis et al., 2007) further showed that the residual activity of the *GAL1*-encoded galactokinase preserves memory in progeny cells by rapidly turning on the Gal4p activator upon cells' re-exposure to galactose.

Despite these studies, a comprehensive example of the role played by neo-Lamarckian epigenetic mechanisms on evolution in the context of a gene network has been lacking. Here, we directly explore the role epigenetic inheritance plays in short-timescale microevolution. We subjected yeast cells to repeated environmental selection based on the expression level of a fluorescent protein reporting on the activity of the canonical GAL network (Figure 1D) (Acar et al., 2005, 2008, 2010; Elison et al., 2018; Luo et al., 2018) over a period of 7 days. We observed reductions in expression level in multiple replicates sorted for the lowest expression that persisted even after the selection pressure was lifted. Using whole-genome sequencing (WGS), chromatin immunoprecipitation (ChIP)-qPCR experiments, and sporulation analysis, we characterized the epigenetic and genetic factors contributing to the persistent expression-level reductions observed.

RESULTS

Applying Environmental Selection on WT GAL Network Activity

To explore the role epigenetic inheritance (Bintu et al., 2016; Bird, 2002; Kouzarides, 2007; Li and Zhang, 2014; Zhou et al., 2011) may play in short-timescale microevolution, we designed an experiment in which a population of isogenic wild-type (WT) yeast cells expressing the yellow fluorescent protein (YFP) under the *GAL1* promoter is subjected to repeated environmental selection in the form of daily sorting based on the expression level of the reporter as measured by flow cytometry. During a 7-day period and corresponding to approximately 101 generations with a 100-min doubling time, the cells were sorted daily based on YFP expression level, and only cells whose expression level is within a particular range (either the lowest 5%, the middle 5%, or the highest 5%; Figure 1E) were selected and allowed to grow further in the same environment (Figures 1B and 1C).

Immediately after sorting, the expression-level distribution of the selected cells is extremely narrow, and it gradually relaxes over time, either to the original distribution if the sorting procedure had no lasting effect on the expression level or to a different distribution with different statistical properties (Figure 1C). By monitoring the expression-level distributions over the 7-day sorting period, one can discern if the sorting intervention had any impact on the expression of the reporter. To determine if any change in the reporter expression observed was transient or lasting, immediately after the 7-day sorting period, the population was grown 3 additional days free from selection pressure

(D) Gene network architecture of the WT yeast GAL network whose activity is reported using a P_{GAL1}-YFP reporter.
(E) Sorting gate determination for WT cells. The off-peak position is determined using expression data measured from uninduced cells (gray bars); the sorting gates in the induced samples (purple bars) are then selected considering only "on" cells.

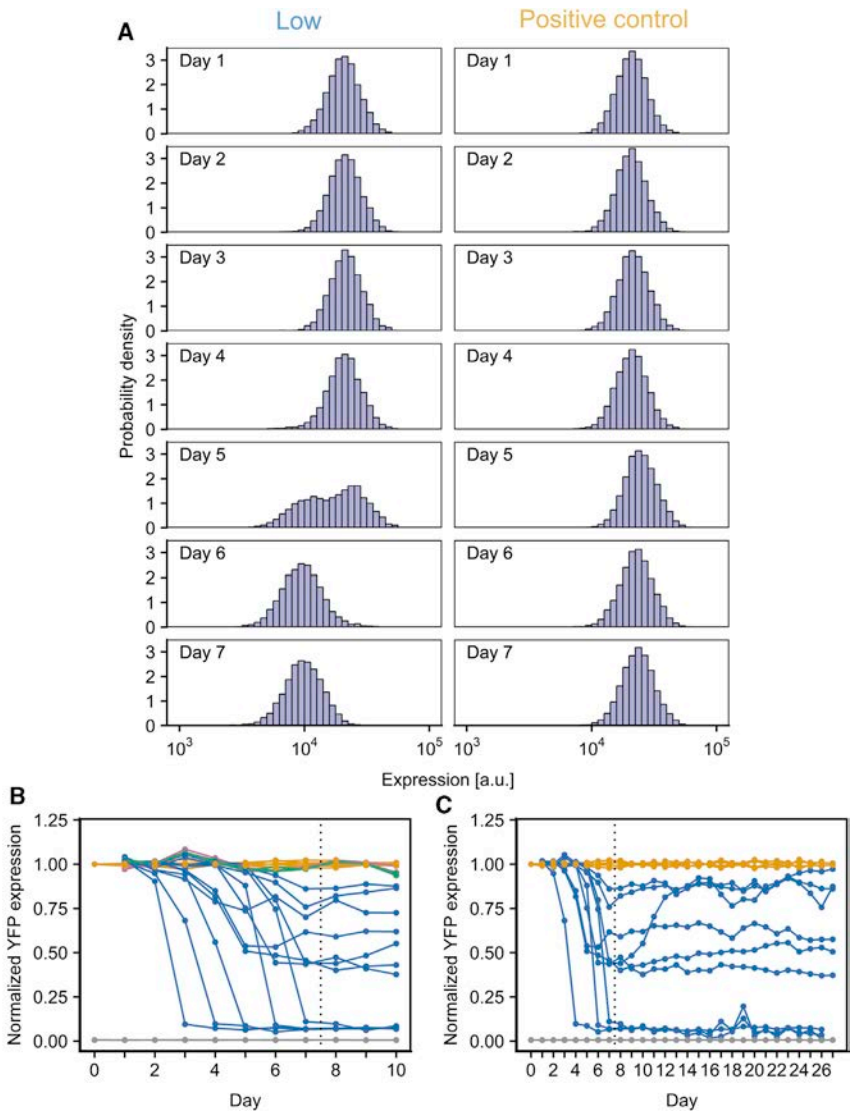


Figure 2. Applying Environmental Selection on WT GAL Network Activity

(A) Sample YFP expression distributions for one of the WT replicates sorted for the lowest expression (left) compared to the positive control (right) over the first 7 days of the experiment.

(B) Normalized YFP expression levels from the YFP-sorting experiment in WT cells. Pink indicates samples sorted for the highest expression, green indicates samples sorted for the median expression, blue indicates samples sorted for the lowest expression, orange indicates positive control, and gray indicates negative control. The dashed line indicates the time at which expression-level-based sorting is terminated. All expression levels are normalized to the corresponding positive control.

(C) Normalized YFP expression levels from the extended YFP-sorting experiment in WT cells. Blue indicates samples sorted for the lowest expression, orange indicates positive control, and gray indicates negative control. The dashed line indicates the time at which expression-level-based sorting is terminated. All expression levels are normalized to the corresponding positive control.

to see if the expression-level distribution reverted back to the original after the selection pressure was lifted.

We found no significant expression-level changes in cells sorted for the middle or highest expression levels (Figure 2B). It is unsurprising that cells already at the middle expression levels retained their character. Given the already high expression levels from the *GAL1* promoter, the lack of change in the cells sorted for the highest expression level may be simply because there is little room for it to increase further. On the other hand, all 12 biological replicates sorted for the lowest expression levels displayed marked reduction in expression (Figures 2A and 2B) to varying degrees that persisted during the 3-day selection-free growth period. Nine replicates were grown (free of selection) for a further 16 days (approximately 230 generations), and 8 of the 9 retained the expression-level reduction (Figure 2C).

To better understand the possible causes driving the observed expression-level reduction, we introduced a second fluorescent reporter protein, mCherry, into the cells. We constructed two strains in which mCherry was either driven by the *TEF1* promoter or the *GAL1* promoter and performed the same YFP-sorting experiment on these strains. We found no change in the expression level of mCherry in the P_{TEF1}-mCherry strain (Figure S1) but significant reduction in the mCherry expression level in the P_{GAL1}-mCherry strain (Figure 3). Further, the level and timing of mCherry expression-level reduction in the P_{GAL1}-mCherry strain was synchronized with that of YFP (Figure 3C). We therefore conclude that the observed reduction in expression was likely due to some factor specific to the GAL network rather than a global factor that can be expected to also affect the expression from P_{TEF1}-mCherry.

Dissecting System Behavior in the Constitutively Active GAL Network

The natural GAL network contains a number of interacting regulators forming feedback loops (Acar et al., 2010; Peng et al., 2016). This complicates the interpretation of the results. For example, the synchronized reduction in double-reporter expression could be due to the dynamics of the epigenetic regulation of the *GAL1* promoter activity, or it could be due to the upstream regulatory elements in the WT GAL network. To eliminate such complication, we deleted the *GAL80* gene—which codes for a repressor through which other GAL network regulators exert their effects—effectively converting the P_{GAL1} promoter into a constitutive promoter dependent only on the transcription factor Gal4p (Figure 4A) and repeated the YFP-sorting experiments in

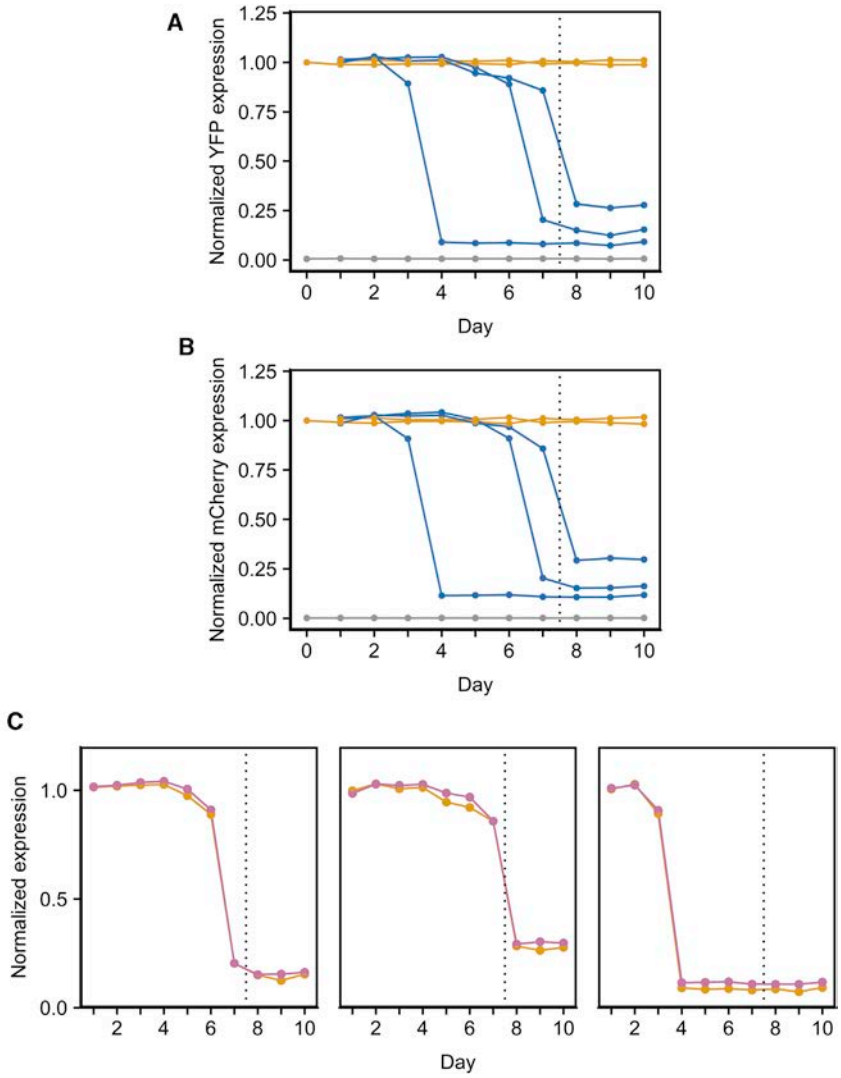


Figure 3. Expression Dynamics from the YFP-Sorting Experiment in WT Cells Containing P_{GAL1}-YFP and P_{GAL1}-mCherry

(A and B) Normalized YFP (A) and mCherry (B) expression levels from the YFP-sorting experiment in WT cells containing P_{GAL1}-YFP and P_{GAL1}-mCherry. Blue indicates samples sorted for the lowest expression, orange indicates positive control, and gray indicates negative control. The dashed line indicates the time at which expression-level-based sorting is terminated. All expression levels are normalized to the positive control.

(C) Comparison of YFP (orange) and mCherry (pink) expression-level trajectories of the three samples sorted for the lowest expression in (A) and (B). Each subpanel represents one sample. The dashed line indicates the time at which expression-level-based sorting is terminated. All expression levels are normalized to the positive control.

See also Figure S1.

this strain. A total of nine biological replicates were used, two of which were found to have accumulated mutations in *GAL4* or the reporter during the course of the experiment and were excluded from further analysis (Data S1).

We found that most biological replicates in which the cells were sorted for the lowest-YFP-expressing cells continued to display a downward shift in mean YFP expression level, although the extent of the shift varied from replicate to replicate, and in one replicate there was no significant change (Figures 4B and 4C, blue curves); five replicates were further grown free of selection for 8 additional days (approximately 115 generations), and all retained their reduced expression levels during this selection-free period (Figures 4E and 4F). Cells sorted for the middle or highest YFP expression did not display any significant changes in either YFP or mCherry expression (Figures 4B and 4C, green and pink curves), just like WT cells. Measuring the doubling times of the cells belonging to colonies isolated from the biological replicate that displayed the largest downward shift in YFP expression (~70%) showed a small increase in doubling times for five colonies isolated from the 7th-day culture, compared to the selection-free positive control (Figure S2A).

We also observed a downward shift in mCherry expression level despite sorting cells in the YFP channel. In all but one of the biological replicates, the YFP and mCherry expression levels were in agreement, but in one replicate, the level of expression-level reduction was significantly different (Figure 4D), though the timing of reduction was similar. This suggests that at least two underlying mechanisms are in play. One mechanism affects the *GAL1* promoter activity generally, driving the synchronized behavior seen both here and previously (Figure 3). But another mechanism, apparently specific to the P_{GAL1}-YFP reporter, must exist that drives the divergence in expression-level reduction between the two reporters, as seen in the last biological replicate.

Measuring Noise Dynamics under Environmental Selection

To see how the selection pressure potentially influences the expression heterogeneity, we next examined the level of noise in P_{GAL1}-YFP and P_{GAL1}-mCherry expression when *gal80Δ* cells were sorted throughout the 7-day period. We discerned no change in the level of P_{GAL1}-YFP noise (coefficient of variation [CV]) in the positive control (no gating) sample or when cells were sorted for the middle or highest expression (Figure S3A).

However, we observed an increase in noise in several samples when cells were sorted for the lowest expression (Figure S3A, blue). Like the change in the expression level itself, this noise increase was stable when the cells were grown selection-free for an additional 10 days after the 7-day sorting period (Figures S3C and S3D, blue). Moreover, in two samples, an accompanying

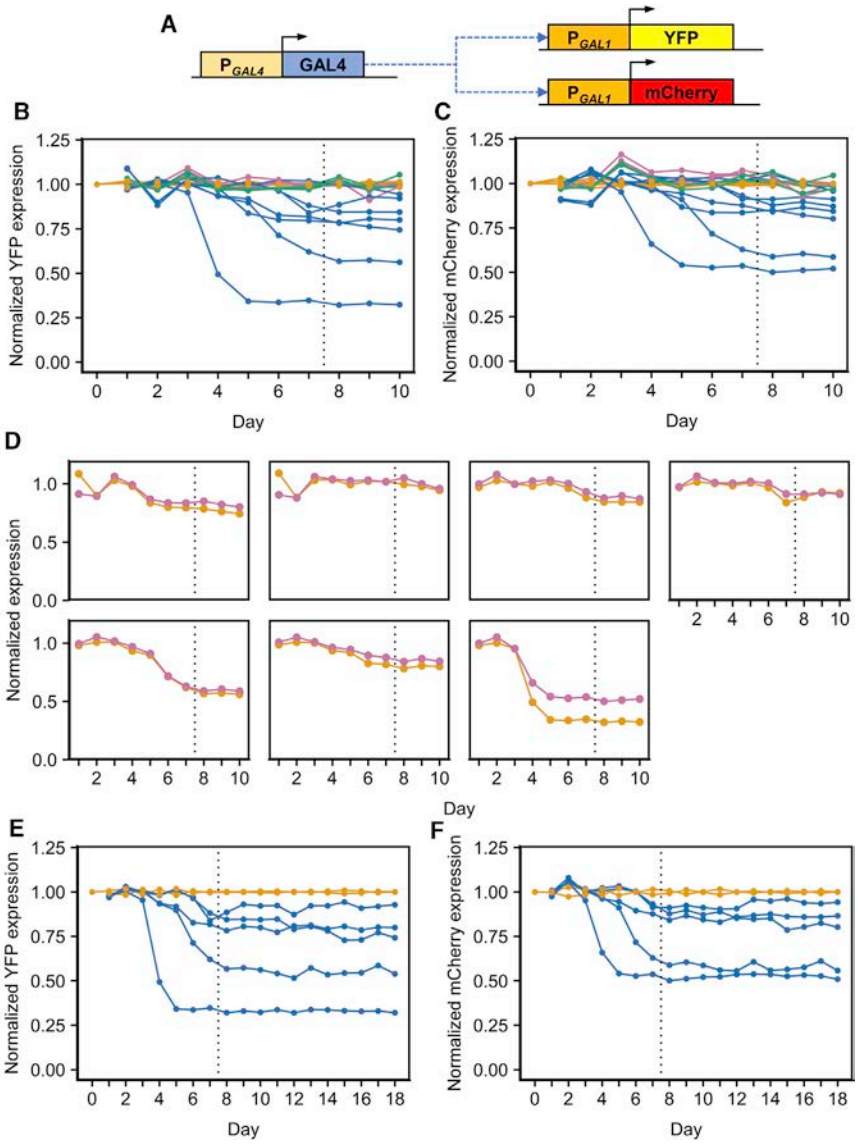


Figure 4. Dissecting System Behavior in the Constitutively Active GAL Network

(A) Gene network architecture of the *gal80Δ* strain.

(B and C) Normalized YFP (B) and mCherry (C) expression levels from the YFP-sorting experiment in *gal80Δ* cells. Pink indicates samples sorted for the highest expression, green indicates samples sorted for the median expression, blue indicates samples sorted for the lowest expression, and orange indicates positive control. The dashed line indicates the time at which expression-level-based sorting is terminated. All expression levels are normalized to the positive control.

(D) Comparison of YFP (orange) and mCherry (pink) expression-level trajectories of the nine samples sorted for the lowest YFP expression in *gal80Δ* cells. Each subpanel represents one sample. The dashed line indicates the time at which expression-level-based sorting is terminated. All expression levels are normalized to the positive control.

(E and F) Normalized YFP (E) and mCherry (F) expression levels from the extended YFP-sorting experiment in *gal80Δ* cells. Blue indicates samples sorted for the lowest expression, and orange indicates positive control. The dashed line indicates the time at which expression-level-based sorting is terminated. All expression levels are normalized to the positive control.

See also Figures S2–S5 and Data S1, S2, S3, and S4.

increase in P_{GAL1}-mCherry noise can be seen when cells were sorted for the lowest P_{GAL1}-YFP expression (Figures S3A–S3D). Under the extreme selection pressure applied during the sorting process, it is unsurprising that higher levels of noise in the protein whose expression is under selection (YFP) may prove evolutionarily beneficial (by increasing the number of cells having reduced expression levels and hence selected during the sorting process). On the other hand, it is likely not advantageous to have higher noise in the expression level of mCherry—or, as a proxy, the structural genes of the *GAL* network, which are responsible for metabolizing the GAL taken from the static environment. Diverging from the optimal level of *GAL* network expression in the environment carries a fitness cost (and higher noise means that more cells are diverging from the optimal level), which could explain why the noise level does not display the same degree of synchronization behavior as the expression-level reduction.

The Effect of the Local Chromatin Environment on Observed Results

To understand the potential influence of genomic loci on the level of expression-level reduction, we performed the same sorting experiment but based on the expression level of P_{GAL1}-mCherry (which was integrated into the *ura3* locus) rather than P_{GAL1}-YFP (which was in the *ho* locus). A total of nine biological replicates were used, two of which were found to have accumulated mutations in *GAL4* or the reporter during the experiment and were excluded from further analysis. Measuring the resulting *GAL1* promoter activity levels during the 7-day sorting period as before, we found expression-level reduction to be significantly more difficult to achieve (Figures 5A–5C), if not impossible, compared to sorting when the reporter cassette is integrated into the *ho* locus, with only one biological replicate out of seven displaying a significant reduction in expression compared to the positive control. We similarly quantified the level of noise in P_{GAL1}-YFP and P_{GAL1}-mCherry expression during these mCherry-sorting experiments (Figures 5D and 5E). We did not detect substantial and persistent changes in noise levels of the kind we had seen previously (Figures S3A and S3B).

Together, these observations suggest that the genomic locus at which the cassette is integrated, and hence the local chromatin structure and epigenetic markers, has a significant effect

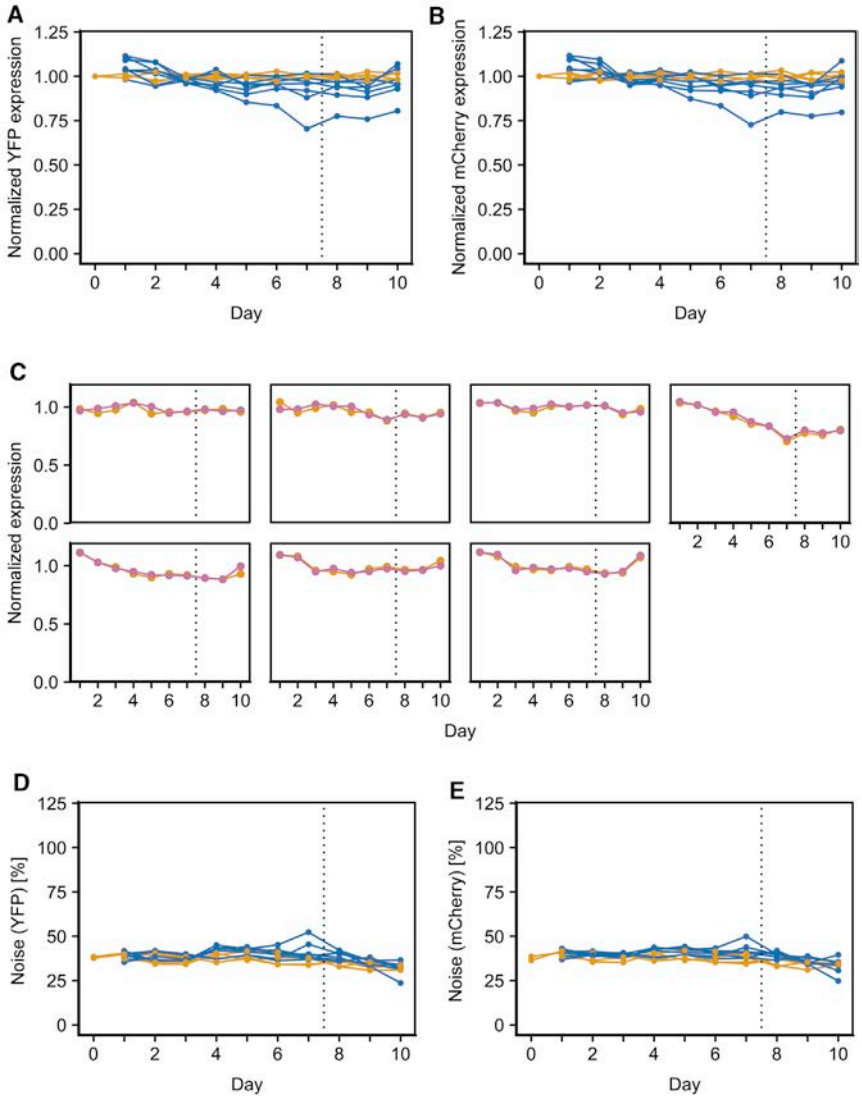


Figure 5. The Effect of the Local Chromatin Environment on Observed Results

(A and B) Normalized YFP (A) and mCherry (B) expression levels from the mCherry-sorting experiment in *gal80Δ* cells. Blue indicates samples sorted for the lowest expression, and orange indicates positive control. The dashed line indicates the time at which expression-level-based sorting is terminated. All expression levels are normalized to the positive control.

(C) Comparison of YFP (orange) and mCherry (pink) expression-level trajectories of the nine samples sorted for the lowest mCherry expression in *gal80Δ* cells. Each subpanel represents one sample. The dashed line indicates the time at which expression-level-based sorting is terminated. All expression levels are normalized to the positive control.

(D and E) Noise in YFP (D) and mCherry (E) expression levels from the mCherry-sorting experiment in *gal80Δ* cells. Blue indicates samples sorted for the lowest expression, and orange indicates positive control. The dashed line indicates the time at which expression-level-based sorting is terminated.

on the phenotype we observed. The *ho* locus appears to be significantly more susceptible to experiencing expression-level reduction in response to the selection pressure we applied compared to the *ura3* locus. Especially given that the genetic mutation rate appears to be approximately constant between the two experiments (in both cases, two out of nine biological replicates were found to have accumulated mutations in the relevant genes), the diverging outcomes strongly suggests that a non-genetic mechanism is involved in suppressing the YFP expression at the *ho* locus.

As noted, sequencing detected no mutations in *GAL4* or the two reporter cassettes (totaling approximately 8 kbp) in the biological replicates under consideration. In addition, we did not detect any substantial fitness changes in the sorted populations passed from one day to the next; if anything, the sorted populations divide slightly slower than unsorted cells (Figure S2). We therefore hypothesized that the observed downward shifts in expression level were due to epigenetic changes in the transcription factor gene *GAL4* and/or in the reporter promoters: the accumulated epigenetic changes "lock" the chromatin into a closed state and are enriched by the daily sorting process. Given the experimental observations, such locks were then necessarily strong enough to persist through hundreds of generations of selection-free growth.

WGS to Explore Genetic Causes of Observed Phenotypes

Next, we performed WGS to evaluate any contributions from global genetic factors on the observed reduction in YFP expression levels. For this, we focused on two biological replicates—the FL6 and FL9 populations in the *gal80Δ* background (STAR Methods)—from which we had seen significant reduction in P_{GAL1}-YFP and P_{GAL1}-mCherry expression on Day7, compared to Day0 expression levels, after gated sorting in the YFP channel. Our local sequencing of the Day7 FL6 and FL9 populations in the reporter cassettes and the *GAL4* region did not identify any mutations.

We isolated five single colonies from the FL6 and FL9 populations on Day7 and randomly selected two single colonies from each of the two groups for performing WGS on them. As controls, we also included in these WGS characterizations two randomly selected single colonies isolated from the Day0 population, as well as two randomly selected single colonies isolated from the positive control group on Day7. Results obtained from the sequencing of each isolated colony were compared to those obtained from the sequencing of the Day0 colonies (Data S2, S3, and S4). While mutations in intergenic promoter regions may also

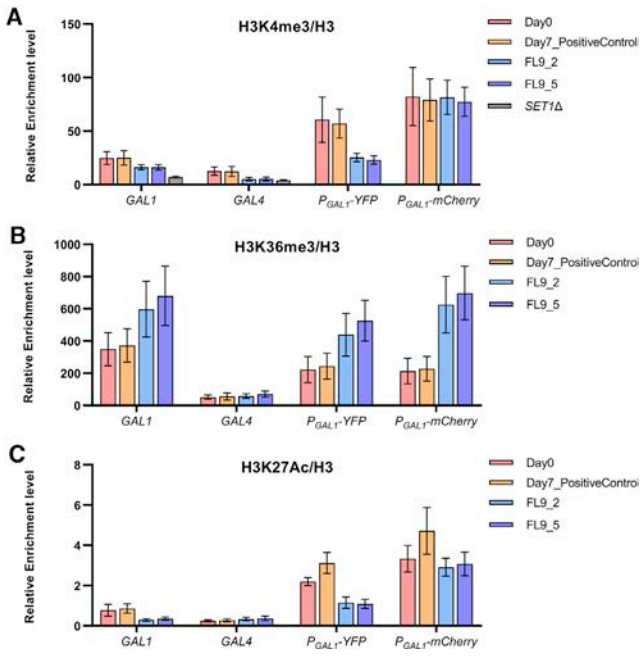


Figure 6. Epigenetic Modification Levels Quantified by ChIP-qPCR

Three types of epigenetic marks—H3K4me3 (A), H3K36me3 (B), and H3K27Ac (C)—were characterized at four genetic loci—*GAL1*, *GAL4*, P_{GAL1}-YFP, and P_{GAL1}-mCherry—in four isogenic populations: Day0, Day7 positive control, FL9_2, and FL9_5. An additional strain, *SET1*Δ, was included for H3K4me3 (A) as a technical control. Error bars indicate SEM (n = 4).

See also Figures S6 and S7 and Tables S1 and S2.

have phenotypic consequences, for the sake of interpretability, we focused on the SNPs (single-nucleotide polymorphisms) causing amino acid alterations in open reading frames (ORFs). We then selected the common ORF mutations found in both single colonies isolated from the FL6 and FL9 populations. Five common mutations were identified for the FL6 colonies (in *FEN2*, *GPM2*, *IRA2*, *NUP133*, and *RPN4*), while the FL9 colonies shared three mutations (in *APL1*, *BDS1*, and *SRB8*).

To see the isolated effects of these mutations on the P_{GAL1}-YFP and P_{GAL1}-mCherry expression levels, we attempted to clone them singly and in combination into a single colony isolated from the unevolved Day0 population using CRISPR. For the FL6 group mutations, we successfully cloned the mutations in the *FEN2*, *IRA2*, and *NUP133* genes one at a time and combinatorically, but the mutations in the *GPM2* and *RPN4* genes could not be cloned because of challenges associated with CRISPR. Measuring the P_{GAL1}-YFP and P_{GAL1}-mCherry expression levels in each constructed strain, we did not see any major changes in expression levels caused by the mutations in *FEN2*, *IRA2*, or *NUP133* genes relative to the Day0 isogenic background without these mutations (Figure S4A). Regarding the mutations in the *GPM2* and *RPN4* genes that could not be cloned, *GPM2* is a nonfunctional homolog of *GPM1* phosphoglycerate mutase, and *RPN4* codes for a transcription factor that stimulates expression of proteasome genes. Despite the potential relevance of *RPN4* for the phenotypes we observed, the degradation experiments we performed for the colonies isolated from the FL6 population did not show differential degradation dynamics for YFP or mCherry (Figures S5A and S5B). Nevertheless, we cannot fully exclude the possibility that the mutations in *GPM2* and/or *RPN4* might exert effects on the phenotypes we observed if they were cloned into the Day0 unevolved, single-colony-derived population.

For the mutations on the *APL1*, *BDS1*, and *SRB8* genes of the FL9 group, on the other hand, we combinatorically constructed all eight strains carrying these mutations one at a time and together. Measuring the P_{GAL1}-YFP and P_{GAL1}-mCherry expression levels in each constructed strain, we saw consistent changes in P_{GAL1}-YFP and P_{GAL1}-mCherry expression levels, relative to the Day0 isogenic background without the mutations (Figure S4B). The degradation experiments we performed for the colonies isolated from the FL9 population did not show differential degradation dynamics for YFP or mCherry (Figures S5C and S5D). The mutation in *SRB8*, coding for a subunit of the RNA polymerase II (RNA Pol II) mediator complex, led to 75% and 50% reductions in the P_{GAL1}-YFP and P_{GAL1}-mCherry expression levels, respectively, across all strains carrying the mutation. Since these expression reductions in levels of YFP and mCherry are very similar to the ones observed at the end of the 7-day sorting period, the mutation in the *SRB8* gene can account for the phenotypic changes observed in one of the biological replicates (sorting group FL9). However, the differential effects of the *SRB8* mutation on YFP and mCherry expression levels indicate that integration-locus-specific epigenetic factors still play a role on the main phenotype of gene expression reduction under environmental selection.

Measuring Acetylation and Methylation Levels on System Components

To further investigate the effect of epigenetic factors on the difference in expression-level decrease between the two reporters as a result of the *SRB8* mutation, we examined the chromatin modification levels at the *GAL4*, *GAL1*, P_{GAL1}-YFP, and P_{GAL1}-mCherry loci in the two WGS-characterized isogenic colonies of the FL9 group (FL9_2 and FL9_5), as well as isogenic colonies isolated from the Day0 population and Day7 positive control population. For this, we tested for three different types of histone modifications via ChIP-qPCR: trimethylation of histone H3 lysine 4 (H3K4me3), which positively correlates with transcriptional activity; trimethylation of histone H3 lysine 36 (H3K36me3), which represses transcription and is known to be associated with HDAC (Histone DeACetylase) recruitment; and acetylation of histone H3 lysine 27 (H3K27ac), which is associated with active transcription.

As expected based on the YFP and mCherry expression levels, ChIP-qPCR results from the Day0 and Day7 positive control colonies showed higher overall H3K4me3 and H3K27Ac levels but lower H3K36me3 levels, compared to the results observed from the FL9_2 and FL9_5 colonies (Figures 6A–6C; Table S1). More specifically, at the endogenous *GAL1* locus, we saw reductions in H3K4me3 and H3K27Ac levels and an increase in H3K36me3 level in the FL9_2 and FL9_5 colonies, compared to the Day7 positive control colony, suggesting that the local transcriptional activities at the *GAL1* locus in FL9_2 and FL9_5 are lower than in Day0 and Day7 positive control. This is consistent with other observations: in FL9_2 and FL9_5,

the YFP protein level was reduced significantly (approximately 70% reduction compared to the Day7 positive control), and there was a significant decrease in the mRNA level of YFP as quantified by qRT-PCR (Figures S6 and S7). We also observed a comparable trend of change in the epigenetic marks at the P_{GAL1}-YFP reporter, indicating that the three types of chromatin modifications tested are similar on these two genetic elements and that both loci are likely governed by the same epigenetic modification machineries that act on the *GAL1* promoter.

We did not see a clear difference between the FL9 colonies and the Day0 and Day7 control colonies at the *GAL4* locus with respect to the H3K36me3 and H3K27Ac modifications, although there seemed to be some reduction of the H3K4me3 mark in FL9_2 and FL9_5. We interpret this as a consequence of a lack of local chromatin-repressing machinery, considering that *GAL4* is expressed constitutively.

Interestingly, while we observed similar trends in H3K36me3 and H3K27Ac modifications at P_{GAL1}-mCherry compared to *GAL1* and P_{GAL1}-YFP, we saw no differentiating trend in the H3K4me3 modification among the tested colonies at P_{GAL1}-mCherry. This divergence suggests a difference in the local chromatin dynamics between the two reporters. It is possible that the locus where P_{GAL1}-mCherry is placed, *URA3* on chromosome V, has a distinct local chromatin regulatory mechanism that interferes with *GAL1*-promoter-specific regulation. The lack of the H3K4me3 mark at P_{GAL1}-mCherry relative to P_{GAL1}-YFP suggests that transcriptional activity at the former is higher than the latter. Indeed, while YFP protein level in FL9_2 and FL9_5 on Day7 was reduced by 70% relative to Day0, mCherry protein level was reduced by only 50%. Therefore, the difference in expression reduction between the two reporters is associable with the difference in the local chromatin modification levels between the two loci where they are located. Together, these results solidify the role of epigenetic modifications on the expression levels of the two reporters.

Sporulation-Based Assessment of Genetic versus Epigenetic Contributions on the Observed Phenotypes

To rule out the possibility that the observed phenotypes of the evolved strains could be explained on purely genetic grounds, we crossed our evolved strains (FL6_2 and FL9_2) with the equivalent of our unevolved WT strain (Day0) of opposite mating type. As a control, we crossed two unevolved WT strains of opposite mating type. After mating, sporulation, and tetrad dissection, we measured the YFP and mCherry expression levels displayed by the progeny of each cross after growing the cells in the same media conditions as used during our evolution experiments. As expected, all offspring of the WT-to-WT cross showed expression levels very similar to the parental strain (Figure 7A).

The FL6_2-to-WT cross generated offspring that was very heterogeneous in expression, contrary to what would be expected from plausible Mendelian genetic mechanisms (Figure 7B). Moreover, the lack of offspring clustering around the parental FL6_2 strain suggested that the mutations on *GPM2* and *RPN4* were not relevant to the observed phenotypes; were they relevant, Mendelian genetics would predict that half (or a quarter) of the offspring spores would carry the mutations on one (or both) of those genes and display a similar phenotype to that of the parental strain, but the fraction of the spores displaying a similar phenotype to that of the parental strain was actually much lower. Surprisingly, a substantial fraction of the spores displayed reporter expression levels higher than the WT parental strain's expression, especially for the YFP reporter. While there was some correlation (Pearson's r = 0.697) between the YFP and mCherry expression levels, we saw that some spores displayed YFP levels higher than the WT but mCherry levels closer to the level displayed by the evolved FL6_2's low mCherry expression level. Meiosis and sporulation are complex cellular programs involving the creation and repair of double-strand breaks on the DNA in certain recombination hotspots (Kolodkin et al., 1986; Lichten and Goldman, 1995; Smith and Nicolas, 1998), and it is known that epigenetic state influences the meiotic recombination hotspots (Brachet et al., 2012). Therefore, a potential explanation for these expression levels is that epigenetic changes carried by the evolved parental strain might have caused unusual meiotic recombination events in the offspring.

The offspring generated from the FL9_2-to-WT cross displayed three clearly distinct clusters of YFP and mCherry expression: one coincided with the parental WT strain's reporter expressions, one coincided with the parental FL9_2 strain's expressions, and the third cluster displayed a WT-like YFP expression and a FL9_2-like mCherry expression (Figure 7C). The reporter expression composition of the third cluster would not be expected based on a purely genetic inheritance pattern, as all genetic components controlling the network are the same for both reporters. Since the only difference between the two genetic constructs (P_{GAL1}-YFP and P_{GAL1}-mCherry) is their chromosomal integration site, we attribute the differential gene expression to the differential impact of the epigenetic marks between the YFP and the mCherry loci, as we described in the previous section.

To explain these three expression clusters displayed by the offspring of the FL9_2-to-WT cross, we propose a model in which the mutation in the RNA Pol II mediator subunit Srb8 contributes to the reporter's downregulation, but epigenetic marks at the YFP reporter locus make its expression level independent of which *SRB8* allele the cell is carrying (Figure 7D), potentially through an epigenetically facilitated compensation mechanism maintaining the overall progression of RNA Pol II irrespective of the Srb8 subunit activity. The inheritance pattern we observed from the offspring of the FL9_2-to-WT crossing supports the presence of epigenetic modifications leading to WT-like YFP expression in one of the three expression clusters. More specifically, the parental FL9_2 strain, which bears both the *SRB8* mutated allele and differential epigenetic marks at the YFP insertion locus (Figures 6A and 6C), displays low YFP and mCherry expression (Figure 7C); the unevolved strain bearing the mutated *SRB8* allele in a WT-like chromatin environment also displays low YFP and mCherry expression (Figure S4B). Therefore, the epigenetic mark we are proposing to explain the offspring's third cluster should favor WT-like YFP expression, as only this inheritance model would generate a 3:1 inheritance pattern for the YFP expression and a 2:2 pattern for the mCherry expression, which matches our observations (Figures 7E and 7F).

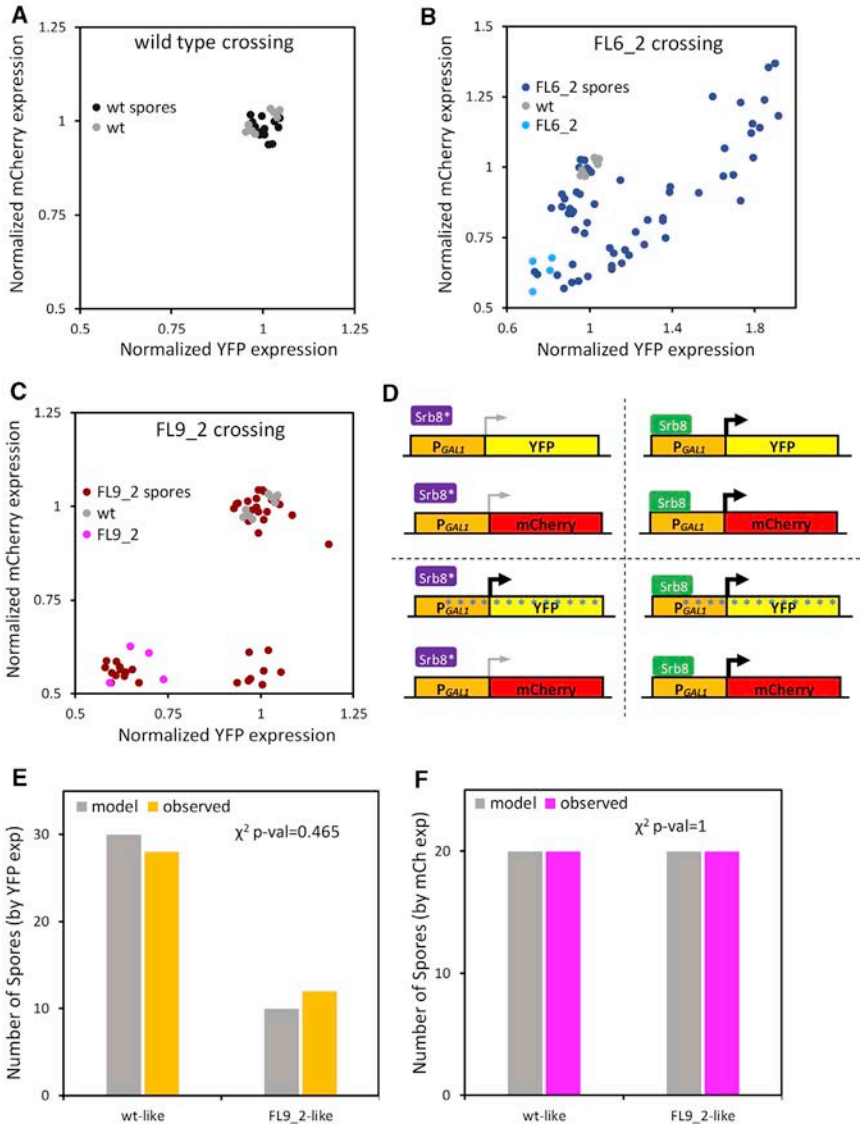


Figure 7. Sporulation-Based Assessment of Genetic versus Epigenetic Contributions on the Observed Phenotypes

(A–C) Scatterplot of the mean measured YFP and mCherry expression displayed by the spores coming from the WT-to-WT crossing (A), the FL6_2-to-WT crossing (B), and the FL9_2-to-WT crossing (C). The expression levels displayed by the parental strains are also shown.

(D) Proposed model explaining the expression distributions of the offspring obtained from the FL9_2-to-WT crossing. The purple Srb8* represents the mutated version identified by WGS, while the green Srb8 represents the WT version. The asterisks along the YFP locus represent the proposed epigenetic mark modulating reporter expression independently from the *SRB8* allele inherited. The color and width of the arrows indicate the strength of the gene expression (thin and gray, FL9_2-like low expression; thick and black, WT-like expression).

(E and F) Bar plots showing the number of spores displaying a given phenotype (WT-like or FL9_2-like), expected by the model and observed experimentally, for the YFP (E) and mCherry (F) expression phenotypes.

Overall, these results indicate that the mechanism behind the observed expression reduction during our evolution experiments cannot be explained only by genetic causes. Epigenetic modifications must be contributing to the differential expression pattern exhibited by the two reporters within the same cell and to the overall evolved phenotype emerging at the end of the evolution process.

DISCUSSION

In this report, we explored the role epigenetic inheritance plays in short-timescale microevolution. We observed reductions in expression level in multiple replicates sorted for the lowest expression that persisted for hundreds of generations, long after the selection pressure was lifted. The amount of decrease in expression level was locus specific, implicating the involvement of local chromatin environment in the process. Performing WGS characterizations on isogenic colonies obtained from two populations, we found that one case of the persistent expression-level reduction was due to genetic factors, while experiments performed for the other case did not indicate a genetic contributor. Measuring the level of chromatin modification marks on system components supported the conclusion that epigenetic regulation differences between integration loci could explain differential YFP and mCherry expression under the same promoter. Finally, results from mating and sporulation experiments provided evidence for the involvement of non-genetic inheritance mechanisms as contributors to the differential expression pattern exhibited by the two reporters in the same cell.

For the replicate that is guided mainly genetically, given that a single mutation in the *SRB8* gene is sufficient to reproduce the decreased YFP and mCherry protein levels measured on Day7, a plausible explanation for this could be that the impaired mRNA synthesis machinery led to a loss in mRNA production and subsequently in protein production in the cell. Since Srb8 is involved in global RNA synthesis, one would expect a global reduction in mRNA levels in the cell as well. We indeed saw a significant reduction in the mRNA level of the housekeeping *ACT1* gene (Figure S7A). Moreover, we found that the isogenic colonies isolated from the FL9 group had significantly increased doubling time compared to both Day0 (~3% increase) and Day7 positive controls (~4% increase), revealing a reduced fitness level that is potentially attributable to inefficient mRNA

synthesis. Despite the genetic contributions to the observed phenotypes, we note that crossing and sporulation analysis showed that the inheritance pattern of this strain's traits was not explainable by a solely genetic mechanism.

Our study involves observation of YFP and mCherry expression dynamics with reporter constructs integrated in different genomic loci. We note that genomic loci differ not only in their local chromatin environment, but also in their sequence context (e.g., presence or absence of certain enhancers), which may be a contributor to expression-level differences between the reporters.

Genetic and epigenetic mechanisms do not have to be mutually exclusive. In response to a particular environmental condition, both kinds of mechanisms can play roles and complement each other. Epigenetic mechanisms generally operate at a shorter timescale than genetic mechanisms, allowing a faster response to changing environmental conditions (Bonduriansky et al., 2012; Burggren, 2016). On the other hand, genetic mechanisms operate at a longer timescale but also produce a more permanent response.

To test whether short-term epigenetic inheritance interacts with genetic change, a recent study (Stajic et al., 2019) used an experimental evolution setup in yeast by tuning low, intermediate, and high levels of heritable silencing of a *URA3* reporter under selection. The authors showed that heritable gene expression through epigenetic chromatin states contributed to adaptive evolution; however, their results and interpretations were not free from mutational effects. More specifically, heritable silencing drove population size expansion and rapid epigenetic adaptation, eventually leading to genetic assimilation of the silent phenotype by mutations. Also, at intermediate or low levels of heritable silencing, the study showed that populations evolved more rapidly by accumulation of adaptive mutations.

Natural environments are usually not fully static, but they fluctuate over time. Memory of previous expression levels, from whatever source, can function as a double-edged sword in fluctuating environments. On one hand, having some memory of the optimal expression level in the current environment confers a fitness advantage in the present (Brickner et al., 2007; Zacharioudakis et al., 2007). On the other hand, locking the expression at a particular level would prevent the cell from responding to environmental changes. Thus, fully persistent memory of expression level would be expected to be detrimental in a fluctuating external environment (Acar et al., 2008; Bódi et al., 2017).

Of the various kinds of heritable factors, genetic mutations are certainly among the most persistent; the cell is full of mechanisms aiming to ensure that genetic materials are faithfully passed from one generation to the next, and moreover, reverting a genetic mutation naturally is even more difficult given the randomness of the mutagenesis process and the rarity of gain-of-function mutations. However, epigenetic mechanisms of reducing gene expression are likely easier to revert when the environment demands it (Klironomos et al., 2013); after all, chromatin is routinely remodeled during the cell cycle (Deniz et al., 2016; Raynaud et al., 2014). Thus, compared to genetic mechanisms, the wide variety of epigenetic mechanisms of regulating gene expression levels are much more easily tuned to environmental demands. While some epigenetic changes may disappear within a few generations (Kundu et al., 2007), others can be highly persistent (Catania et al., 2020). Thus, the epigenetic toolset allows the cell to strike a balance between memorizing gene expression states and being plastic to external environmental changes.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - *Saccharomyces* cerevisiae with the W303 genetic background
- METHOD DETAILS
  - Construction of yeast strains and plasmids
  - Flow cytometry data acquisition and sorting
  - mRNA transcript levels determination by qRT-PCR
  - Local sequencing of key GAL network components
  - Whole Genome Sequencing (WGS) sample preparation
  - Measuring doubling-times of cell populations
  - SNP introduction with CRISPR-Cas9
  - Degradation dynamics of fluorescent proteins
  - ChIP-qPCR experiments
  - Mating, sporulation and tetrad dissection
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Flow cytometry data analysis
  - Whole Genome Sequencing data analysis
  - χ^2 test for checking inheritance models for sporulation outcomes

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2020.108306.

ACKNOWLEDGMENTS

We thank J. Gendron, E. Karatekin, L. Mirny, G.P. Wagner, and Acar Lab members for comments and feedback on various stages of this work. X.L. acknowledges support through the China Scholarship Council-Yale World Scholars Program. M.H. acknowledges funding from the US National Institutes of Health (GM136325). M.A. acknowledges funding from the US National Institutes of Health (1DP2AG050461-01 and 1U54CA209992-01).

AUTHOR CONTRIBUTIONS

X.L., R.S., and M.A. designed the experiments and analyses, interpreted the data and results, and designed and prepared the manuscript. X.L. constructed the strains, performed the experiments, collected the data, and contributed to data analysis. R.S. analyzed the data and performed statistical analyses. X.L., H.-Y.R., and M.H. performed the ChIP-qPCR experiments and helped interpret their results. D.F.M. performed the mating and sporulation experiments, including strain construction; flow-cytometry measurements and analyses; and interpreting, plotting, and writing the results with M.A. D.F.M. also

contributed to the revision of the manuscript. M.H. dissected the tetrads. M.A. conceived and supervised the project. All authors read and approved the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 22, 2019
Revised: August 17, 2019
Accepted: October 2, 2020
Published: October 27, 2020

REFERENCES

Acar, M., Becskei, A., and van Oudenaarden, A. (2005). Enhancement of cellular memory by reducing stochastic transitions. Nature *435*, 228–232.

Acar, M., Mettetal, J.T., and van Oudenaarden, A. (2008). Stochastic switching as a survival strategy in fluctuating environments. Nat. Genet. *40*, 471–475.

Acar, M., Pando, B.F., Arnold, F.H., Elowitz, M.B., and van Oudenaarden, A. (2010). A general mechanism for network-dosage compensation in gene circuits. Science *329*, 1656–1660.

Ahn, S.H., Kim, M., and Buratowski, S. (2004). Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3′ end processing. Mol. Cell *13*, 67–76.

Avery, O.T., Macleod, C.M., and McCarty, M. (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types: Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type iii. J. Exp. Med. *79*, 137–158.

Barbieri, E.M., Muir, P., Akhuetie-Oni, B.O., Yellman, C.M., and Isaacs, F.J. (2017). Precise Editing at DNA Replication Forks Enables Multiplex Genome Engineering in Eukaryotes. Cell *171*, 1453–1467.e13.

Bintu, L., Yong, J., Antebi, Y.E., McCue, K., Kazuki, Y., Uno, N., Oshimura, M., and Elowitz, M.B. (2016). Dynamics of epigenetic regulation at the single-cell level. Science *351*, 720–724.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. Genes Dev. *16*, 6–21.

Bódi, Z., Farkas, Z., Nevozhay, D., Kalapis, D., Lázár, V., Csörgő, B., Nyerges, Á., Szamecz, B., Fekete, G., Papp, B., et al. (2017). Phenotypic heterogeneity promotes adaptive evolution. PLoS Biol. *15*, 1–26.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Bonduriansky, R., and Day, T. (2009). Nongenetic Inheritance and Its Evolutionary Implications. Annu. Rev. Ecol. Evol. Syst. *40*, 103–125.

Bonduriansky, R., Crean, A.J., and Day, T. (2012). The implications of nongenetic inheritance for evolution in changing environments. Evol. Appl. *5*, 192–201.

Brachet, E., Sommermeyer, V., and Borde, V. (2012). Interplay between modifications of chromatin and meiotic recombination hotspots. Biol. Cell *104*, 51–69.

Brickner, D.G., Cajigas, I., Fondufe-Mittendorf, Y., Ahmed, S., Lee, P.C., Widom, J., and Brickner, J.H. (2007). H2A.Z-mediated localization of genes at the nuclear periphery confers epigenetic memory of previous transcriptional state. PLoS Biol. *5*, e81.

Burggren, W.W. (2014). Epigenetics as a source of variation in comparative animal physiology - or - Lamarck is lookin' pretty good these days. J. Exp. Biol. *217*, 682–689.

Burggren, W. (2016). Epigenetic Inheritance and Its Role in Evolutionary Biology: Re-Evaluation and New Perspectives. Biology (Basel) *5*, 24.

Burkhardt, R.W., Jr. (2013). Lamarck, evolution, and the inheritance of acquired characters. Genetics *194*, 793–805.

Catania, S., Dumesic, P.A., Pimentel, H., Nasif, A., Stoddard, C.I., Burke, J.E., Diedrich, J.K., Cook, S., Shea, T., Geinger, E., et al. (2020). Evolutionary Persistence of DNA Methylation for Millions of Years after Ancient Loss of a De Novo Methyltransferase. Cell *180*, 263–277.e20.

Chatterjee, M., and Acar, M. (2018). Heritable stress response dynamics revealed by single-cell genealogy. Sci. Adv. *4*, e1701775.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al.; 1000 Genomes Project Analysis Group (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158.

Darwin, C. (1859). On the origin of species by means of natural selection, or, the preservation of favoured races in the struggle for life (J. Murray).

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

Deniz, Ö., Flores, O., Aldea, M., Soler-López, M., and Orozco, M. (2016). Nucleosome architecture throughout the cell cycle. Sci. Rep. *6*, 19729.

Elison, G.L., Xue, Y., Song, R., and Acar, M. (2018). Insights into Bidirectional Gene Expression Control Using the Canonical GAL1/GAL10 Promoter. Cell Rep. *25*, 737–748.e4.

Giaever, G., Chu, A.M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., et al. (2002). Functional profiling of the Saccharomyces cerevisiae genome. Nature *418*, 387–391.

Griffiths, A.J.F., Miller, J.H., Suzuki, D.T., Lewontin, R.C., and Gelbart, W.M. (2000). An Introduction to Genetic Analysis, Seventh Edition (W. H. Freeman).

Hahne, F., LeMeur, N., Brinkman, R.R., Ellis, B., Haaland, P., Sarkar, D., Spidlen, J., Strain, E., and Gentleman, R. (2009). flowCore: a Bioconductor package for high throughput flow cytometry. BMC Bioinformatics *10*, 106.

Halfmann, R., Jarosz, D.F., Jones, S.K., Chang, A., Lancaster, A.K., and Lindquist, S. (2012). Prions are a common mechanism for phenotypic inheritance in wild yeasts. Nature *482*, 363–368.

Huang, S. (2009). Non-genetic heterogeneity of cells in development: more than just noise. Development *136*, 3853–3862.

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

Janke, C., Magiera, M.M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., Moreno-Borchart, A., Doenges, G., Schwob, E., Schiebel, E., and Knop, M. (2004). A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. Yeast *21*, 947–962.

Kaiser, C., Michaelis, S., and Mitchell, A. (1994). Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual (Cold Spring Harbor Laboratory Press).

Kaufmann, B.B., Yang, Q., Mettetal, J.T., and van Oudenaarden, A. (2007). Heritable stochastic switching revealed by single-cell genealogy. PLoS Biol. *5*, e239.

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

Kolodkin, A.L., Klar, A.J.S., and Stahl, F.W. (1986). Double-strand breaks can initiate meiotic recombination in S. cerevisiae. Cell *46*, 733–740.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693–705.

Kronholm, I., and Collins, S. (2016). Epigenetic mutations can both help and hinder adaptive evolution. Mol. Ecol. *25*, 1856–1868.

Kundu, S., Horn, P.J., and Peterson, C.L. (2007). SWI/SNF is required for transcriptional memory at the yeast GAL gene cluster. Genes Dev. *21*, 997–1004.

Kuzawa, C.W., and Thayer, Z.M. (2011). Timescales of human adaptation: the role of epigenetic processes. Epigenomics *3*, 221–234.

Lacal, I., and Ventura, R. (2018). Epigenetic Inheritance: Concepts, Mechanisms and Perspectives. Front. Mol. Neurosci. *11*, 292.

Laland, K., Uller, T., Feldman, M., Sterelny, K., Müller, G.B., Moczek, A., Jablonka, E., Odling-Smee, J., Wray, G.A., Hoekstra, H.E., et al. (2014). Does evolutionary theory need a rethink? Nature *514*, 161–164.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv, arXiv:1303.3997. https://arxiv.org/abs/1303.3997.

Li, E., and Zhang, Y. (2014). DNA methylation in mammals. Cold Spring Harb. Perspect. Biol. *6*, a019133.

Lichten, M., and Goldman, A.S.H. (1995). Meiotic recombination hotspots. Annu. Rev. Genet. *29*, 423–444.

Luo, X., Song, R., and Acar, M. (2018). Multi-component gene network design as a survival strategy in diverse environments. BMC Syst. Biol. *12*, 85.

McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., Flicek, P., and Cunningham, F. (2016). The Ensembl Variant Effect Predictor. Genome Biol. *17*, 122.

Nei, M., and Nozawa, M. (2011). Roles of mutation and selection in speciation: from Hugo de Vries to the modern genomic era. Genome Biol. Evol. *3*, 812–829.

Olson-Manning, C.F., Wagner, M.R., and Mitchell-Olds, T. (2012). Adaptive evolution: evaluating empirical support for theoretical predictions. Nat. Rev. Genet. *13*, 867–877.

Peng, W., Liu, P., Xue, Y., and Acar, M. (2015). Evolution of gene network activity by tuning the strength of negative-feedback regulation. Nat. Commun. *6*, 6226.

Peng, W., Song, R., and Acar, M. (2016). Noise reduction facilitated by dosage compensation in gene networks. Nat. Commun. *7*, 12959.

Raynaud, C., Mallory, A.C., Latrasse, D., Jégu, T., Bruggeman, Q., Delarue, M., Bergounioux, C., and Benhamed, M. (2014). Chromatin meets the cell cycle. J. Exp. Bot. *65*, 2677–2689.

Ryu, H.-Y., and Ahn, S. (2014). Yeast histone H3 lysine 4 demethylase Jhd2 regulates mitotic rDNA condensation. BMC Biol. *12*, 75.

Sikorski, R.S., and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics *122*, 19–27.

Skinner, M.K. (2015). Environmental epigenetics and a unified theory of the molecular aspects of evolution: A neo-Lamarckian concept that facilitates neo-Darwinian evolution. Genome Biol. Evol. *7*, 1296–1302.

Skinner, M.K., Guerrero-Bosagna, C., and Haque, M.M. (2015). Environmentally induced epigenetic transgenerational inheritance of sperm epimutations promote genetic mutations. Epigenetics *10*, 762–771.

Smith, K.N., and Nicolas, A. (1998). Recombination at work for meiosis. Curr. Opin. Genet. Dev. *8*, 200–211.

Stajic, D., Perfeito, L., and Jansen, L.E.T. (2019). Epigenetic gene silencing alters the mechanisms and rate of evolutionary adaptation. Nat. Ecol. Evol. *3*, 491–498.

Tyedmers, J., Madariaga, M.L., and Lindquist, S. (2008). Prion switching in response to environmental stress. PLoS Biol. *6*, e294.

Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., et al. (2013). From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. Curr. Protoc. Bioinformatics *43*, 11.10.1–11.10.33.

Wach, A., Brachat, A., Pöhlmann, R., and Philippsen, P. (1994). New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. Yeast *10*, 1793–1808.

Xue, Y., and Acar, M. (2018a). Mechanisms for the epigenetic inheritance of stress response in single cells. Curr. Genet. *64*, 1221–1228.

Xue, Y., and Acar, M. (2018b). Live-Cell Imaging of Chromatin Condensation Dynamics by CRISPR. iScience *4*, 216–235.

Zacharioudakis, I., Gligoris, T., and Tzamarias, D. (2007). A yeast catabolic enzyme controls transcriptional memory. Curr. Biol. *17*, 2041–2046.

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

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
α-H3K4me3	Abcam	Cat# ab8580; RRID:AB_306649
α-H3K36me3	Abcam	Cat# ab9050; RRID:AB_306966
α-H3Ac	Millipore	Cat# 07-360; RRID:AB_310550
α-H3	Abcam	Cat# ab1791; RRID:AB_302613
Experimental Models: Organisms/Strains		
S. cerevisiae: Strain background: W303 MATα	Our Lab Stocks	MA0001
S. cerevisiae: Strain background: W303 MATa	Our Lab Stocks	MA0002
S. cerevisiae: Strain background: W303 MATα ho::HIS5-P_{GAL1}-YFP	Our Lab Stocks	WP35
S. cerevisiae: Strain background: W303 MATα, ho::HIS5-P_{GAL1}-YFP, ura3::URA3-P_{GAL1}-mCherry	This paper	WP35URAPg1mC
S. cerevisiae: Strain background: W303 MATα, ho::HIS5-P_{GAL1}-YFP, ura3::URA3-P_{TEF1}-mCherry	This paper	WP35URAPtefmC
S. cerevisiae: Strain background: W303 MATα, ho::HIS5-P_{GAL1}-YFP, ura3::URA3-P_{GAL1}-mCherry, gal80Δ::NatNT2	This paper	XLUYmCgΔ80
S. cerevisiae: Strain background: W303 MATa, ho::HIS5-P_{GAL1}-YFP, ura3::URA3-P_{GAL1}-mCherry, gal80Δ::KanMX4	This paper	DMY375
Recombinant DNA		
Plasmid: pRS306	(Sikorski and Hieter, 1989)	N/A
Plasmid: pRS305	(Sikorski and Hieter, 1989)	N/A
Plasmid: pYM17	(Janke et al., 2004)	N/A
Plasmid: pFA6-kanMX4	(Wach et al., 1994)	N/A
Plasmid: HIS5-P_{GAL1}-YFP	Our Lab Stocks	N/A
Plasmid: URA3-P_{GAL1}-mCherry	This paper	N/A
Plasmid: URA3-P_{TEF1}-mCherry	This paper	N/A
Plasmid: pRS314-CAS9	Our Lab Stocks	N/A
Chemicals, Peptides, and Recombinant Proteins		
Gibson Assembly® Master Mix	New England BioLabs	E2611S
Restriction Enzyme: BstBI	New England BioLabs	R0519S
iTaqTM Universal SYBR® Green Supermix	Bio-Rad	1725120
Protein G-Sepharose	GE Healthcare	17-0618-01
Pronase	Roche	11 459 643 001
High Capacity RNA-to-cDNA kit	Applied Biosystems	4388950
YeaStar Genomic DNA Kit	ZYMO Research	D2002
Cycloheximide	Sigma-Aldrich	C7698
cOmplete protease inhibitor cocktail	Roche	11697498001
PMSF	AmericanBio	AB01620
β-glucuronidase	Sigma-Aldrich	G7017

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Oligonucleotides		
Primers for qRT-PCR and ChIP-qPCR, see Table S2	This paper	N/A
Deposited Data		
GenBank: SAMN11440943	GenBank	N/A
GenBank: SAMN11440944	GenBank	N/A
GenBank: SAMN11440945	GenBank	N/A
GenBank: SAMN11440946	GenBank	N/A
GenBank: SAMN11440947	GenBank	N/A
GenBank: SAMN11440948	GenBank	N/A
GenBank: SAMN11440949	GenBank	N/A
GenBank: SAMN11440950	GenBank	N/A
Software and Algorithms		
NEBuilder® Assembly Tool	New England BioLabs	N/A
BD FACSuite	BD Biosciences	N/A
R	www.R-project.org	N/A
Bioconductor flowCore package	(Hahne et al., 2009)	N/A
Trimmomatic	(Bolger et al., 2014)	N/A
BWA-MEM	(Li, 2013)	N/A
Picard's MarkDuplicates	https://github.com/broadinstitute/picard	N/A
GATK tools	(Van der Auwera et al., 2013)	N/A
VCFtools	(Danecek et al., 2011)	N/A
VEP tool	(McLaren et al., 2016)	N/A
Other		
BD FACS-Aria	Becton Dickinson	N/A
Cary 60 UV-Vis Spectrometer	Agilent Technologies	N/A
Zeiss Tetrad "Advanced Yeast Dissection Microscope"	Carl Zeiss	N/A
Illumina HiSeq4000	Illumina	N/A

RESOURCE AVAILABILITY

Lead Contact

Requests for further information and for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Murat Acar (murat.acar@yale.edu).

Materials Availability

Yeast strains and plasmids used in this study are described in the Key Resources Table and will be made available upon request from the Lead Contact, Murat Acar (murat.acar@yale.edu).

Data and Code Availability

The accession numbers for the sequencing data from WGS runs reported in this paper are GenBank: SAMN11440943, GenBank: SAMN11440944, GenBank: SAMN11440945, GenBank: SAMN11440946, GenBank: SAMN11440947, GenBank: SAMN11440948, GenBank: SAMN11440949, GenBank: SAMN11440950. These numbers are also listed in the Key Resources Table.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Saccharomyces cerevisiae with the W303 genetic background

All yeast *Saccharomyces* cerevisiae strains constructed are based on the haploid W303 strain background. Complete genotypic descriptions of all strains can be found in the Key Resources Table.

All cultures were grown in synthetic minimal media with histidine dropout and appropriate supplements of other amino-acids. Culture growths were performed in a 30°C shaker (225rpm) in a volume of 1mL. After 48hrs of growth on histidine-dropout minimal media plates containing 2% glucose, strains were grown in liquid minimal media for 22hr ("overnight") in the presence of 0.1% mannose as a non-inducing sole carbon source. This was followed by a 72hr induction period in liquid minimal media containing 0.1% mannose and 0.2% galactose as carbon sources.

METHOD DETAILS

Construction of yeast strains and plasmids

Strains used to study the GAL network were built on WP35 which is a haploid wild-type strain carrying single copy of the P_{GAL1}-*YFP* reporter in the *ho* locus. The double reporter strains carrying a second reporter (P_{GAL1}-*mCherry-tCYC1* or P_{TEF1}-*mCherry-tCYC1*) were constructed with the following steps. First, plasmids carrying P_{GAL1}-*mCherry-tCYC1* or P_{TEF1}-*mCherry-tCYC1* on the pRS306 backbone (Sikorski and Hieter, 1989) were constructed using the Gibson Assembly® Master Mix and NEBuilder® Assembly Tool (New England BioLabs). The resulting plasmids were then linearized within the *URA3* gene at BstBI cut site and transformed into WP35 using the standard lithium acetate (LiOAc) transformation technique. qPCR was performed to select colonies carrying single copy of the second reporter. To obtain the strain XLUYmCgΔ80, the P_{AgTEF}-*natNT2-tADH1* cassette from pYM17 (Euroscarf, Janke et al., 2004) was integrated into the double reporter strain carrying P_{GAL1}-*YFP-tCYC1* and P_{GAL1}-*mCherry-tCYC1* to replace the *GAL80* gene by using 60bps homology regions immediately before and after *GAL80*.

For the mating and sporulation experiments, a MATa counterpart for XLUYmCgΔ80 strain was constructed, starting with the MA0002 strain. First, a single copy of *HIS5*-P_{GAL1}-*YFP-tCYC1* was inserted into the *ho* locus by using 60bp homology regions around the *ho* locus. Then, the P_{GAL1}-*mCherry-tCYC1* plasmid was linearized by BstBI and inserted in single-copy into the *URA3* locus. Finally, the *GAL80* ORF was deleted by amplifying the P_{AgTEF}-*KanR-tAgTEF1* cassette from the *pFA6-kanMX4* plasmid (Wach et al., 1994) with 60bp homology regions immediately before and after the *GAL80* ORF.

Flow cytometry data acquisition and sorting

After the induction period described in the "Experimental Model and Subject Details" section above, the expression distribution of ~50,000 cells were measured by flow cytometry (FACS-Aria; Becton Dickinson) at flow rates between 4 to 8 (flow rate scale of 1-11 corresponds to approximately 10-80 μL/min), and the cell-sorting period was initiated. During the 7-day-long sorting period, populations underwent expression-based sorting once a day, followed by the selection-free growth period lasting from 3 to 20 days during which the entire expression distribution was passed from one day to the next instead of gated-sorting. Every 24hrs during the 7-day-long gated-sorting period, individual cells were sorted into fresh media of the same type by applying fluorescence intensity based gates rendering the highest, middle and lowest 4.8%–5.2% of the total cell population (referred to as HIGH, MID, LOW). In "forward" sorting groups, the gates were selected based on YFP fluorescence, while in the "reverse" sorting groups they were selected based on mCherry fluorescence. 450 individual cells were sorted for the HIGH groups, and 600 cells were sorted for all other groups. To minimize potential variations in the size and/or morphology of the sorted cells, the gating process also involved applying a narrow FSC-SSC (ForwardSCatter-SideSCatter) range corresponding to the densest ~20% of the total cell population. Grown cultures taken from Day0, Day7, the last day of the selection-free period, as well as certain other days throughout each sorting period were frozen on the same day for further analysis. Starting from the overnight growth period and until the end of the gated and selection-free sorting periods, cell densities were kept low (between OD_{600} 0.2 and 0.3) to prevent nutrient depletion.

mRNA transcript levels determination by qRT-PCR

Selected cell populations of the strain XLUYmCgΔ80 frozen after sorting on Day0 and Day7 were recovered as single isogenic colonies from glycerol stocks streaked on 2% glucose minimal media plates with histidine dropout. Colonies were then grown overnight in liquid minimal media containing 0.1% mannose as the sole carbon source, and induced in minimal media containing 0.1% mannose and 0.2% galactose for 48hrs in 50mL volume, reaching a final OD_{600} of less than 0.2. Fluorescence levels of the induced cells were recorded by flow cytometry right before harvesting for total RNA. cDNA was prepared by using the High Capacity RNA-to-cDNA kit from Applied Biosystems. The resulting cDNA was then used in qPCR reactions to quantify mRNA levels of genes of interest. For qPCR, we used the iTaq™ Universal SYBR® Green Supermix from Bio-Rad and targeted 4 genes with 2 sets of primers for each: ACT1 (primer pair EPIACT1-2F and EPIACT1-2R were used as the endogenous control for ΔC_T calculation), YFP, mCherry, GAL1, and GAL4. The relative transcription levels for samples within the same sorting experimental group were calculated with the Day0 population's transcript levels used as the control. The qPCR primers used are listed on Table S2. The amplicons were between 158 and 161bps long.

Local sequencing of key GAL network components

To see whether or not mutations were accumulated on the genetic components relevant to the GAL network activity, frozen cell populations from Day0 and Day7 sorting groups of the strain XLUYmCgΔ80 (the groups that showed at least 20% decrease in YFP or mCherry expression compared to corresponding Day0 expression) were recovered from glycerol stocks streaked on 2% glucose minimal media plates with histidine dropout. Populations were then grown overnight in liquid minimal media containing 0.1%

mannose as the sole carbon source, and induced in minimal media containing 0.1% mannose and 0.2% galactose for 48hrs in 1mL volume, reaching a final OD_{600} of less than 0.3. After the induction period, expression measurements were performed by flow cytometry and also genomic DNA contents were extracted from the induced populations to sequence key genetic components of the GAL network. All re-induced Day7 populations exhibited similar expression levels relative to expression levels of their corresponding Day0 populations; in other words, freezing and re-induction after the actual sorting process did not alter the relative expression levels in these populations. For sequencing, we selected the LOW ("L") sorting groups of the strain XLUYmCgΔ80 which showed over 20% decrease in reporter expression; the local sequencing was performed for the P_{GAL1}-YFP, P_{GAL1}-mCherry, and P_{GAL4}-GAL4 constructs from the beginning of the promoters to the end of the terminators. With "F or R" indicating "Forward or Reverse" sorting based on YFP or mCherry, these Day7 sequenced groups of sorted populations were named as FL1, FL2, FL6, FL8, FL9, RL2, RL5, and RL6; the numbers indicate the identity of the biological replicate from the 7-day-long sorting experiment. Sequencing was first performed on the population level for these sorting groups; genomic DNA was prepared from the entire sorted populations from Day7, and no apparent mutation was identified. Then, 5 randomly-selected single colonies were isolated from each population for isogenic expression characterization, and sequencing was performed on select single colonies which had similar expression profile as the corresponding original population. No mutation on the P_{GAL1}-YFP, P_{GAL1}-mCherry, P_{GAL4}-GAL4 constructs was found in colonies isolated from the FL1, FL6, FL9, RL2, RL5, whose full sequences are given in Data S1; however, mutations/changes were found in colonies isolated from FL2, FL8 and RL6 (Data S1).

Whole Genome Sequencing (WGS) sample preparation

Out of the 5 single colonies isolated from the FL6 and FL9 populations on Day7 (for which local-sequenced for the P_{GAL1}-YFP, P_{GAL1}-mCherry, P_{GAL4}-GAL4 constructs did not identify any mutations), we randomly selected 2 single colonies from each of the FL6 and FL9 groups for performing whole genome sequencing on them. As controls, we also included in these WGS characterizations 2 randomly-selected single colonies isolated from the Day0 population, as well as 2 randomly selected single colony isolated from the positive control group on Day7.

Cells from each single colony were recovered from glycerol stock on 2% glucose minimal media plates with histidine dropout, and grown in 10mL YPD liquid media until the cell-density (OD_{600}) reached ~1. The YeaStar Genomic DNA Kit (ZYMO Research) was used for genomic DNA extraction. The process was repeated 2-3 times until 1-5 μg of purified DNA ($OD_{260/280}$ between 1.8 and 2) concentrated in 50 μL of TE buffer was acquired for each sample. The purified DNA were pooled and sequenced at the Yale Center for Genome Analysis with Illumina HiSeq4000 (paired-end, 150bp) targeting 200X coverage.

Measuring doubling-times of cell populations

Five isogenic colonies isolated from each of the Day7 FL6 and FL9 populations, as well as one isogenic colony isolated from each of the Day0 and Day7 positive controls, were recovered from glycerol stocks and streaked on 2% glucose minimal media plates with histidine dropout. Colonies were then grown overnight in liquid minimal media containing 0.1% mannose as the sole carbon source, and induced in minimal media containing 0.1% mannose and 0.2% galactose for 48hrs in 1mL volume, reaching a final OD_{600} of less than 0.3. Following the induction period, cultures were continuously grown in the same media conditions, and the growth rate analyses were performed based on the dilution rates and the OD_{600} values measured at 6 different time points across the next 52-55hrs. At each time point, all cultures were diluted to maintain OD_{600} below 0.55 to keep growth at log-phase and to prevent nutrition depletion. The average log-phase doubling-time $t_{doubling}$ was calculated (Figures S2A and S2B) using the following formula:

$$t_{doubling} = t_{duration} \Big/ \log_2\left(\frac{D_{end}}{D_{start}} * \prod_{k=1}^{N} d_k\right)$$

$t_{duration}$: duration between the start and end of the continuous culture growth
D_{end}: OD_{600} at the end of continuous culture growth
D_{start}: OD_{600} at the start of continuous culture growth
N: total number of dilutions (here N = 6)
d_k: dilution rate at time point k

SNP introduction with CRISPR-Cas9

Select mutations identified from WGS were cloned into a single colony from Day0 to see if each or all of them could result in the phenotypic changes observed in the sorting experiment. To choose candidate mutations, we first selected mutations within ORFs that cause changes in their corresponding amino acids. We then selected the common mutations found in both single colonies isolated from a group: 5 common mutations for FL6 (*FEN2*, *GPM2*, *IRA2*, *NUP133, RPN4*), and 3 common mutations for FL9 (*APL1*, *BDS1, SRB8*). To introduce these mutations back into a single colony from Day0, a centromeric plasmid with backbone pRS314 carrying a Cas9 cassette was first transformed into a Day0 single colony. For each mutation, the resulting strain was transformed with

donor DNA carrying the mutation together with a plasmid with backbone pRS305 carrying guide RNA cassette targeting the mutation site. The final strains were sequenced locally to verify the intended genetic alterations.

Degradation dynamics of fluorescent proteins

Five isogenic colonies isolated from each of the Day7 FL6 and FL9 populations, as well as one isogenic colony isolated from each of the Day0 and Day7 positive control, were recovered from glycerol stocks and streaked on 2% glucose minimal media plates with histidine dropout. Colonies were then grown overnight in liquid minimal media containing 0.1% mannose as the sole carbon source, and induced in minimal media containing 0.1% mannose and 0.2% galactose for 48hrs, reaching a final OD_{600} of ~0.2 in a total volume of 5mL for each sample. Cycloheximide (Sigma-Aldrich, C7698) was then added to the cultures at the final concentration of 10 μg/mL. Samples were taken from the cycloheximide-treated cultures for fluorescence measurements with flow cytometry (FACS-Aria, Becton Dickinson) at the following time points while the cultures continued being incubated in a 30°C shaker: 0hr (right before adding cycloheximide), 15min (right after adding cycloheximide), 1.5hr, 4hr, 6hr, 21hr and 28.5hr.

During the flow cytometry measurements, due to the cycloheximide treatment, changes were observed in the position of the total cell population based on the FSC-SSC readings. Therefore, a large FSC-SSC gate covering the densest 40%–80% of the total population was applied.

ChIP-qPCR experiments

From the single colonies selected for WGS, we selected the two colonies from FL9 together with one single colony from each of the Day0 and Day7 positive control group for histone modification characterization with ChIP-qPCR. We tested three types of histone modifications at GAL1, GAL4, YFP and mCherry open reading frames: H3K4me3, H3K36me3 and H3Ac. A yeast strain with *set1Δ* background was used as negative control for the H3K4me3 group. Cells from each single colony were recovered from glycerol stock on 2% glucose minimal media plates with histidine dropout, grown overnight in minimal media containing 0.1% mannose as the sole carbon source, and induced in minimal media containing 0.1% mannose and 0.2% galactose for 48hrs. We used 300mL of culture with OD_{600} ~0.4 to initiate ChIP.

ChIP experiments were performed as described previously (Ahn et al., 2004; Ryu and Ahn, 2014). Briefly, formaldehyde was added to a final concentration of 1% for 20 min. Cross-linking was quenched by addition of glycine to 240 mM. Cells were collected by centrifugation, washed in TBS twice, and then lysed with glass beads in FA lysis buffer {50 mM HEPES-KOH at pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na deoxycholate, 0.1% SDS, cOmplete protease inhibitor cocktail (Roche, 11697498001), 1 mM PMSF (AmericanBio, AB01620)}. Sheared chromatin by sonication was incubated with Protein G-Sepharose (GE Healthcare, 17-0618-01) bound with anti-H3K4me3 (Abcam, ab8580), anti-H3K36me3 (Abcam, ab9050), anti-H3Ac (Millipore, 07-360), or anti-H3 (Abcam, ab1791). Following washings, eluted chromatin fragments were treated with pronase (Roche, 11 459 643 001), and DNA was purified by phenol/chloroform extraction. qPCR assays were performed using 1:8 diluted DNA template, and then the results for methylated or acetylated H3 were normalized to total histone H3 signals and the internal control (a fragment amplified from an untranscribed region on ChrIV).

Forward and reverse primer sequences used for ChIP-qPCR are listed on Table S2. EPIYFP pair targets 169-327bp from 5′ of YFP ORF; EPImC-1 pair targets 177-338bp from 5′ of mCherry ORF; EPIGal1-1 pair targets 319-476bp from 5′ of GAL1 ORF; EPIGal4-1 pair targets 2225-2385bp from 5′ of GAL4 ORF; IntIV pair is the endogenous control and targets an intergenic region on chromosome IV.

Mating, sporulation and tetrad dissection

The desired strains were grown in YPD plates overnight. Then, a small amount of the fresh patch of cells was mixed with its mating counterpart in a fresh YPAD plate (YPD supplemented with 20mg/L Ade) and incubated for 4h at 30°C. Zygote formation was checked by microscopy and a portion of the mating patch was transferred to YPD+Nat+G418 plate, which selected for diploid cells. Single diploid colonies were transferred to GNA pre-sporulation plates (Giaever et al., 2002) (5% glucose, 3% nutrient broth, 1% yeast extract, 2% agar) and grown for a day at 30°C, and then transferred to sporulation plates (Kaiser et al., 1994) (1% potassium acetate, 0.1% yeast extract, 0.05% dextrose, 2% agar) where they were incubated at room temperature for 4-5 days. Tetrad dissection was performed on a YPD plate after degrading the ascus wall with β-glucuronidase (Sigma-Aldrich, G7017), and then incubated at 30°C for 2 days. The spores coming from each tetrad were spotted on YPD+G418 and YPD+Nat plates to check that they were displaying a proper segregation pattern of the markers, which qualified them for further analysis.

QUANTIFICATION AND STATISTICAL ANALYSIS

Flow cytometry data analysis

Each sample of flow cytometry (FACSAria, Becton Dickinson) data were analyzed in R using the Bioconductor flowCore software package (Hahne et al., 2009). The FSC-SSC gate was chosen to cover the densest portion of the total population and eliminate individuals with unusual morphologies, such as dying cells and cell debris; the same gate was used for all samples gathered during a single experiment. Each FACS sample had on average ~7500 cells after gating. Log-amplified fluorescence measurements for the

gated cells were converted to linear scale for analysis. When needed for the wild-type strain, a threshold for ON state was selected based on fluorescence measurements from uninduced cells and applied uniformly to all relevant samples.

The raw expression level of each strain on each day during the multi-day experiment is measured by averaging the single-cell reporter fluorescence as measured by flow cytometry. To control for the effect of day-to-day variations, the raw expression levels were normalized using the average expression level of the same reporter in the positive control samples measured on the same day.

Whole Genome Sequencing data analysis

The sequencing reads were first trimmed using Trimmomatic (Bolger et al., 2014) with the following settings: "LEADING:3 TRAILING:3 SLIDINGWINDOW:2:30" (Barbieri et al., 2017). The filtered reads from each sample were independently aligned to the current version of S288C reference genome from Ensembl using BWA-MEM (Li, 2013) and converted to BAM format using SAMtools. Picard's MarkDuplicates (https://github.com/broadinstitute/picard) was used to mark duplicates in the resulting BAM files. We then realigned the reads with the GATK tools (Van der Auwera et al., 2013) RealignerTargetCreator and IndelRealigner, and variants were called using GATK's HaplotypeCaller with "-ploidy 1." SNPs and Indels were extracted from the resulting file and filtered manually using GATK's SelectVariants and VariantFiltration. For SNP, we used the following parameters: -filterExpression "QD<2.0 || FS>50.0 || MQ<50.0 || SOR>3.0 || MQRankSum<-12.5 || ReadPosRankSum<-8.0." For Indel, we used these parameters: -filterExpression "QD<2.0 || FS>200.0 || InbreedingCoeff<-0.8 || SOR>10.0 || ReadPosRankSum<-20.0." Finally, we used VCFtools (Danecek et al., 2011) to identify variants in FL6, FL9 and Day7-positive-control samples relative to Day0, which were then annotated with Ensembl's VEP tool (McLaren et al., 2016).

χ^2 test for checking inheritance models for sporulation outcomes

The proposed inheritance model was tested as described (Griffiths et al., 2000). Briefly, the χ^2 test statistic was calculated as $\sum_{i=1}^{n} \frac{(E_i - O_i)^2}{E_i}$, where n is the number of classes for a phenotype, E_i is the expected number of individuals under that class according to the model, and O_i is the experimentally-observed number of individuals classified under that class. This test statistic was compared with the χ^2 distribution with n-1 degrees of freedom to obtain the likelihood of the experimental observation if the assumed model was true (p value).

EVOLUTIONARY BIOLOGY

Two different epigenetic information channels in wild three-spined sticklebacks are involved in salinity adaptation

Melanie J. Heckwolf[1]*†, Britta S. Meyer[1†‡], Robert Häsler[2], Marc P. Höppner[2], Christophe Eizaguirre[3], Thorsten B. H. Reusch[1]

Epigenetic inheritance has been proposed to contribute to adaptation and acclimation via two information channels: (i) inducible epigenetic marks that enable transgenerational plasticity and (ii) noninducible epigenetic marks resulting from random epimutations shaped by selection. We studied both postulated channels by sequencing methylomes and genomes of Baltic three-spined sticklebacks (*Gasterosteus aculeatus*) along a salinity cline. Wild populations differing in salinity tolerance revealed differential methylation (pop-DMS) at genes enriched for osmoregulatory processes. A two-generation experiment demonstrated that 62% of these pop-DMS were noninducible by salinity manipulation, suggesting that they are the result of either direct selection or associated genomic divergence at cis- or trans-regulatory sites. Two-thirds of the remaining inducible pop-DMS increased in similarity to patterns detected in wild populations from corresponding salinities. The level of similarity accentuated over consecutive generations, indicating a mechanism of transgenerational plasticity. While we can attribute natural DNA methylation patterns to the two information channels, their interplay with genomic variation in salinity adaptation is still unresolved.

INTRODUCTION

Recent advances in epigenetics challenge our understanding of inheritance and adaptive evolution (*1*–*3*). It has been suggested that epigenetic modifications—for example, via DNA methylation, histone modification, or small RNAs—create phenotypic diversity and ultimately contribute to rapid evolutionary adaptation (*4*–*6*). Several theoretical models posit that the heritable proportion of these molecular modifications can be classified into two distinct information channels (*5*, *7*, *8*). Selection-based epigenetic marks emerge as spontaneous epimutations that remain stable across subsequent generations, although their overall stability is three to four orders of magnitude lower compared to DNA base changes (*7*, *9*). Similar to adaptation from DNA sequence–based variation, these epimutations may result in different phenotypes that become targets of natural selection and thereby carry information on past selection regimes without directly responding to the current environment (*5*, *8*, *10*). On the other hand, detection-based effects describe inducible epigenetic marks at defined genomic locations, which are under environmental control (*7*). Such transfer of parental information linked to environmental cues represents a rapid and reliable mechanism underlying transgenerational plasticity, which is hypothesized to buffer the extinction risk of populations under sudden environmental change until genetic adaptations can catch up ("genetic rescue") (*7*, *11*). Distinguishing between these mechanisms is important because they have very different implications for the evolution of populations. Stable epigenetic marks follow evolutionary principles of DNA sequence–based inheritance with random variation shaped by selection. In contrast, directional processes via inducible epigenetic marks can be considered a transgenerational form of plasticity that involve previously evolved regulatory mechanisms targeting specific sites on the genome. While the principal differences between these two transmission channels are clear (*4*, *5*, *7*), empirical evidence for their presence in wild vertebrate populations is lacking.

Here, we assess whether these two epigenetic information channels can be detected in nature and test whether short-term acclimation responses match patterns of DNA methylation variation of locally adapted populations. Transgenerational experiments that yield DNA methylation profiles more similar to those of locally adapted natural populations would provide evidence that DNA methylation is mechanistically involved in adaptive transgenerational plasticity.

Studying adaptation to ocean salinity is particularly suited to identification of selection- and detection-based effects because spatiotemporal patterns in ocean salinity are more stable than other variables, for instance, temperature. Since salinity change imposes strong physiological stress with well-defined cellular effects (*12*), natural salinity gradients offer unparalleled opportunities to use local patterns of epigenetic variation as background against which direction and magnitude of results from experimental salinity manipulations can be tested. One suitable ecosystem to follow such a space-for-time approach is the Baltic Sea, which is a semi-enclosed marginal sea that has been dubbed a "time machine" to evaluate the predicted perturbations associated with global change (*13*).

Taking advantage of the Baltic Sea salinity gradient, we sequenced the methylomes [reduced representation bisulfite sequencing (RRBS)] and whole genomes of three-spined sticklebacks (*Gasterosteus aculeatus*) from three populations that are locally adapted to different salinities [6, 20, and 33 practical salinity units (PSU)] (*14*) in and outside the Baltic Sea. Specifically, we focus on the patterns of (epi)genomic variation, while transgenerational phenotypic effects have been described previously for the exact same populations (*15*).

[1]Marine Evolutionary Ecology, GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany. [2]Institute of Clinical Molecular Biology, Kiel University, Kiel, Germany. [3]School of Biological and Chemical Sciences, Queen Mary University of London, London, UK.
*Corresponding author. Email: mheckwolf@geomar.de
†These authors contributed equally to this work.
‡Present address: Max Planck Research Group Behavioural Genomics, Max Planck Institute for Evolutionary Biology, Plön, Germany.

Baltic stickleback populations are genetically differentiated [genome-wide average pairwise F_{ST} = 0.028 (*14*)] and show patterns consistent with local adaptation to salinity regimes in controlled common garden experiments (*15*, *16*). Moreover, previous studies have revealed transgenerational plasticity in response to variation in temperature (*17*) and changes in DNA methylation levels at osmoregulatory genes in response to within-generational salinity manipulation (*18*, *19*). However, it remains unclear whether DNA methylation mediates transgenerational plasticity, a possible mechanism enabling adaptive phenotypes to rapidly emerge in the face of environmental change. In this study, we consider transgenerational effects to be adaptive if the preacclimation of the parents enhances the fitness of the offspring, sometimes referred to as intergenerational effects. To address this question, we complemented our field survey with a two-generation salinity acclimation experiment using the mid-salinity population (20 PSU). This experiment enabled us to quantify the proportion of noninducible (stable, potentially selection-based) and inducible (potentially detection-based) DNA methylation within and across generations (Fig. 1), acknowledging that we tested methylation mark stability only with respect to experimental salinity manipulation. We focused on the methylation of cytosines at cytosine-phosphate-guanine dinucleotides (CpG sites), the most common methylation motif in vertebrates (*20*), with partial inheritance potentially involved in adaptive evolution (*11*).

We tested three nonexclusive hypotheses: (i) Stickleback populations from different salinities (6, 20, and 33 PSU) show differentially methylated CpG sites (hereafter referred to as pop-DMS). (ii) Such pop-DMS include both types of methylation sites: experimentally stable sites (potentially selection based) and experimentally inducible sites (potentially detection based). (iii) Upon transgenerational salinity acclimation, inducible DNA methylations become more similar to the patterns of natural populations at corresponding salinities. When associated with beneficial phenotypic effects and increased relative fitness, the latter would be evidence for a mechanism underlying adaptive transgenerational plasticity (overview in Fig. 2).

RESULTS

Identifying differentially methylated CpG sites between stickleback populations along a natural salinity cline

pop-DMS were determined via RRBS in 46 wild-caught sticklebacks from three different sites that varied in average salinity [Sylt (SYL), 33 PSU; Kiel (KIE), 20 PSU; Nynäshamn (NYN), 6 PSU; Fig. 1]. After quality and coverage filtering, we obtained 525,985 CpG sites present in all groups ($q < 0.0125$; methylation difference, ≥15%), corresponding to ~4% of all CpG sites in the stickleback genome. Among pairs of wild-caught populations, we detected 1470 (comparison of 20 versus 6 PSU) and 1158 (20 versus 33 PSU) pop-DMS. The distribution of these sites was random with regard to the genomic features (promoter, exon, intron, and intergenic; 20 versus 6 PSU: $X^2_3 = 3.36$, $P = 0.340$; 20 versus 33 PSU: $X^2_3 = 1.61$, $P = 0.656$; table S1) and chromosomal regions (fig. S1A). Among these pop-DMS, 1098 (20 versus 6 PSU) and 871 (20 versus 33 PSU) were located close to [<10 kb from transcription start sites (TSS)] or within genes and thereby associated with 655 and 510 genes, respectively. Many of these genes are involved in fundamental biological processes such as DNA repair and strand renaturation, as well as chromosome condensation and separation (fig. S2). Of particular relevance is the enrichment in genes associated with osmoregulatory processes such as ion transport and channel activity, renal water homeostasis and absorption, and urine volume regulation (Fig. 3).

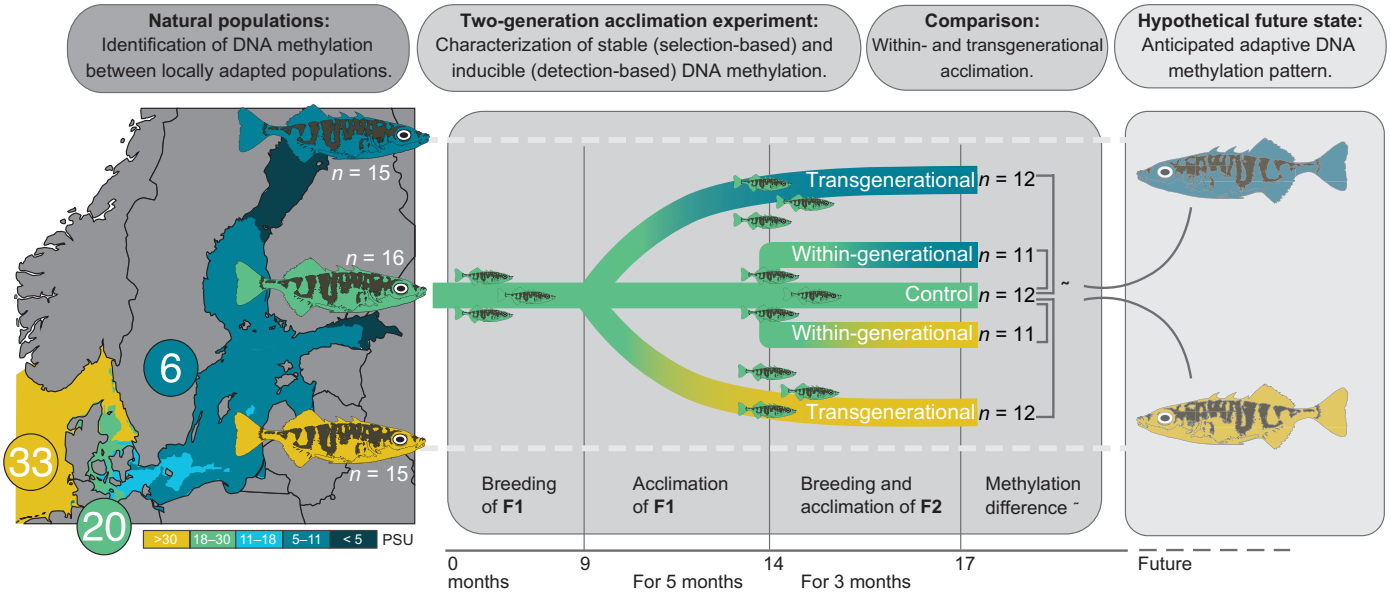


Fig. 1. Experimental space-for-time approach. We characterized DNA methylation profiles (via RRBS) and whole genomes [whole-genome sequencing (WGS)] of fish from three populations of wild-caught three-spined sticklebacks locally adapted to 6 (blue; *n* = 15), 20 (green; *n* = 16), and 33 (yellow; *n* = 15) PSU. We also bred and acclimated sticklebacks from the mid-salinity location (20 PSU) within one ("within-generational") or over two ("transgenerational") generations to decreased (6 PSU) or increased (33 PSU) salinity while maintaining a control group at its original salinity (*n* = 11 to 12 per group; see details in the figure). Differential methylation within and across generations was assessed and compared to natural populations locally adapted to the corresponding salinity, serving as the hypothetical future DNA methylation state to capture long-term adaptation processes.

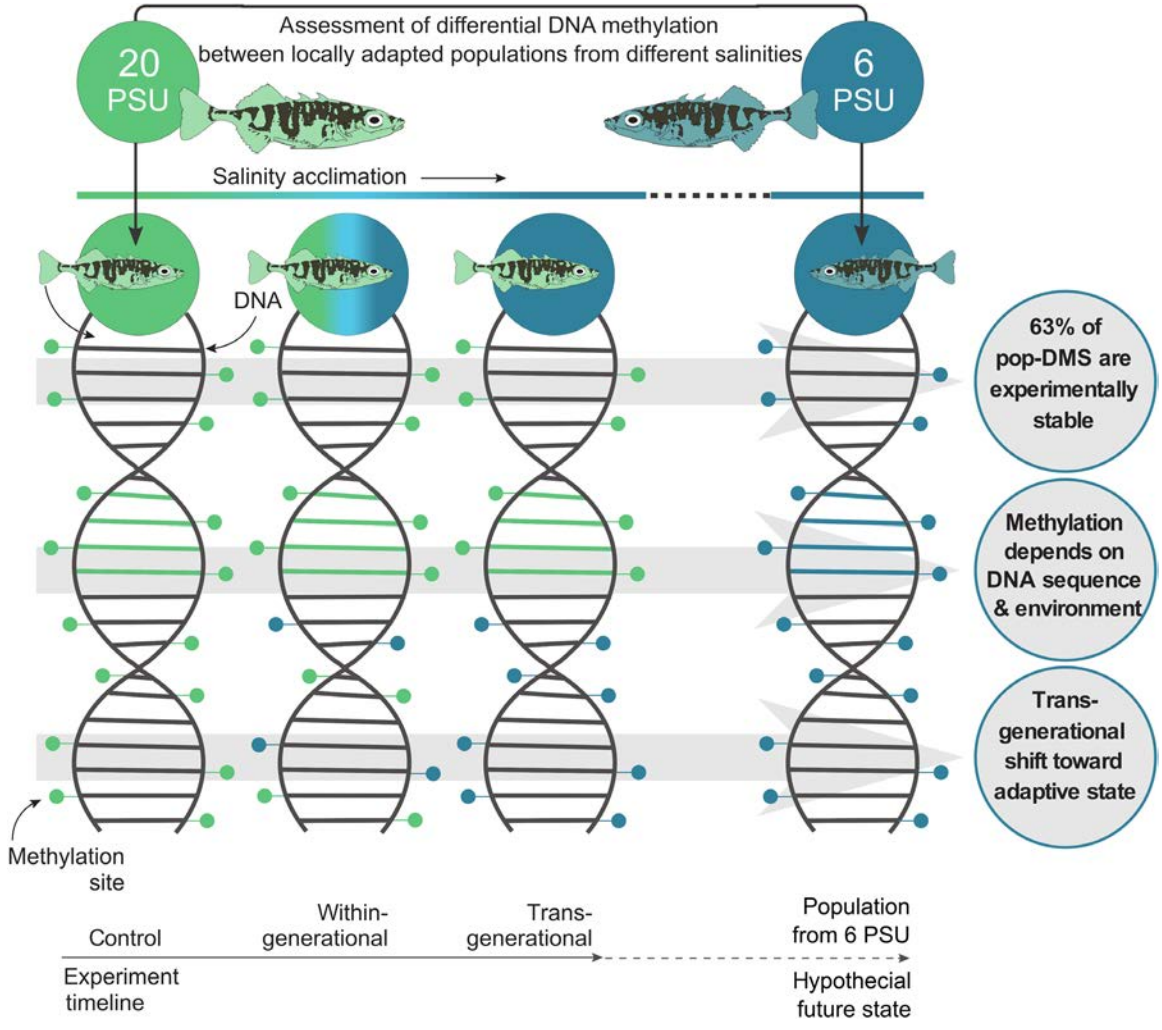


Fig. 2. Graphical summary of the main results. We used the Baltic Sea salinity gradient to study the role of DNA methylation in local salinity adaptation and the response to salinity change in a space-for-time approach. To assess the potential future acclimatization and adaptation processes of the natural stickleback population from 20 PSU (KIE; green) to the predicted desalination (*63*), we compared differences in DNA methylation at CpG sites between wild-caught and laboratory-bred sticklebacks. Following the experiment timeline (bottom), we compared methylation levels of the experimental control group from 20 PSU to within- and trans-generational acclimation of 20 PSU sticklebacks to 6 PSU (DNA from left to right). The population locally adapted to 6 PSU serves as the hypothetical future state in which salinities will decrease (blue; DNA on the right). The three main results are written in the circles with schematically and horizontally corresponding DNA methylation changes. (i) Sixty-three percent of the DMS between the populations remained stable under experimental salinity change. (ii) The direction of experimental methylation change was dependent not only on the treatment but also on the degree of genetic differentiation between the populations [see Fig. 4 (A to D) for results]. (iii) Trans-generational salinity acclimation shifted DNA methylation patterns closer to the anticipated adaptive state found in the hypothetical future population [see Fig. 4 (E to H) for results]. For clarity, only one (6 PSU) of the two foreign salinity regimes tested (6 and 33 PSU) is shown. The results for the experimental fish acclimated to 33 PSU were very similar (see Fig. 1 for full experimental design and Fig. 4 for results).

Genes associated with ≥10 pop-DMS are listed in Table 1 [for all genes, see table S2 (A and B)].

Characterizing stable and inducible DNA methylation in a two-generation salinity acclimation experiment

To assess the proportion of inducible DNA methylation, we conducted a two-generation salinity acclimation experiment with laboratory-bred sticklebacks from the mid-salinity population that was subjected to either increased or decreased salinity (Fig. 1). We considered pop-DMS to be noninducible (hereafter referred to as "stable") when both the within-generational and the transgenerational acclimation groups were not differentially methylated compared to the control group ($q \geq 0.0125$). On the other hand, if a pop-DMS was differentially methylated between at least one of the acclimation groups (within- and transgenerational) compared to the control group ($q < 0.0125$; methylation difference, ≥15%), then this site was considered inducible. Pop-DMS with a significant q value not exceeding the threshold of differential DNA methylation were treated as a separate category (hereafter referred to as inconclusive). After two generations of salinity acclimation, we found that most of the pop-DMS remained stable, regardless of the direction of salinity change (926 pop-DMS, 63% at decreased salinity; 694 pop-DMS, 60% at increased salinity). A smaller number of pop-DMS (13%) were inducible, as they showed a significant change in CpG

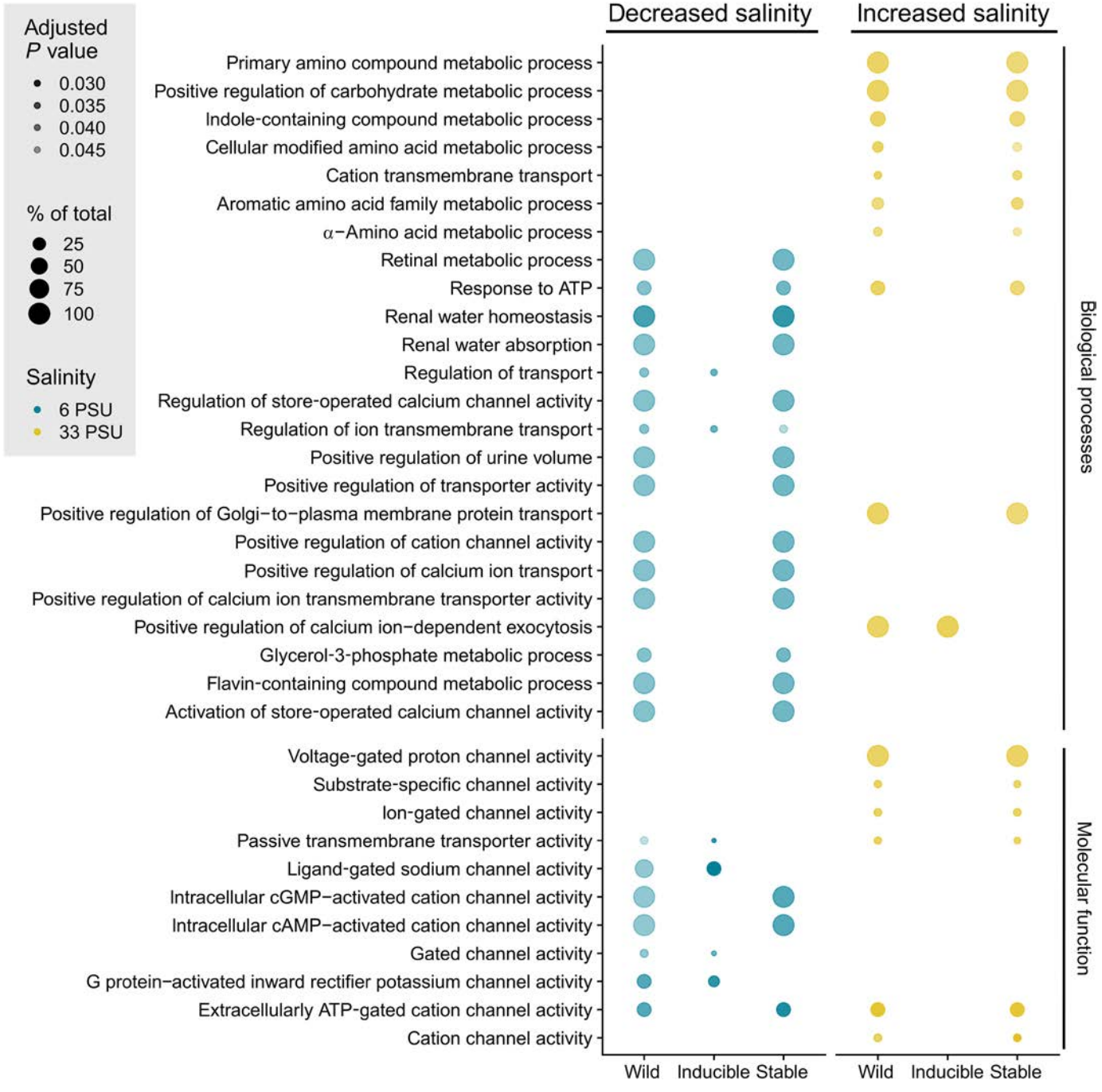


Fig. 3. Gene Ontology terms for biological processes and molecular functions. Gene Ontology (GO) terms for biological processes and molecular functions under salinity increase (20 versus 33 PSU; yellow) and decrease (20 versus 6 PSU; blue) associated with pop-DMS are presented. The graph is split into GO terms associated with pop-DMS from natural stickleback populations across a salinity cline (wild) and their experimental inducibility (inducible and stable) in a two-generation acclimation experiment. The size of the circles refers to the number of genes of this term in the groups (in %), and the transparency refers to the false discovery rate–corrected *P* value (darker circles refer to a lower adjusted *P* value). This subset is filtered for GO terms including the following keywords: "channel," "transport," "water," "chloride," "potassium," "homeostasis," "ion-dependent," "urine," "ATP" (adenosine 5′-triphosphate), and "metabolic"; see fig. S2 for the full figure. cGMP, guanosine 3′,5′-monophosphate; cAMP, adenosine 3′,5′-monophosphate; G protein, heterotrimeric GTP-binding protein.

methylation upon experimental salinity decrease (198 pop-DMS) or increase (148 pop-DMS). An additional 24 and 27% (346 and 316 pop-DMS, respectively) were inconclusive. The number of inducible pop-DMS (13%) derived from comparisons between natural populations was much higher than expected from a random subset of CpG sites across the genome (<1%; 1000 replicates; salinity decrease: $X^2_2 = 1090.7$, $P < 0.001$; salinity increase: $X^2_2 = 967.7$, $P < 0.001$). This means that pop-DMS are enriched for sites that plastically

Table 1. Differentially methylated genes across natural populations along a salinity cline. Genes derived from DNA methylation comparisons between natural populations associated with ≥10 pop-DMS [decreased salinity: KIE (20 PSU) versus NYN (6 PSU); increased salinity: KIE (20 PSU) versus SYL (33 PSU)]. Ensembl gene ID and name as well as the position on the chromosome are listed. The numbers refer to the numbers of DMS in the population comparison (wild). These DMS were classified into inducible, inconclusive, and stable sites according to their behavior in a two-generation salinity acclimation experiment with laboratory-bred sticklebacks from the mid-salinity population (20 PSU) exposed to experimental salinity increase or decrease (33 and 6 PSU, respectively). Furthermore, inducible sites were distinguished whether they matched methylation levels of the locally adapted population (expected) or not (opposite). Genes written in bold vary in both population comparisons. We used a Fisher's exact test to assess whether pop-DMS associated to the same gene are correlated in their response to experimental salinity change (nonrandom distribution among the categories stable, inducible, and inconclusive) and reported corresponding *P* values. For a full table on all genes associated with one or more pop-DMS, see table S2 (A and B).

Ensembl gene ID	Chromosome	Start position	End position	Gene name	Wild	Inducible	Expected inducible	Opposite inducible	Stable	Inconclusive	Fisher's exact (*P*)
Salinity decrease:											
ENSGACG00000008328	Chr10	12860144	12863850	*si:dkey-166k12.1*	24	0	0	0	9	15	0.005
ENSGACG00000019416	**Chr7**	**4451892**	**4453656**	***HMX1 ortholog***	**17**	**0**	**0**	**0**	**9**	**8**	**0.033**
ENSGACG00000013229	**Chr18**	**15327717**	**15352321**		**15**	**0**	**0**	**0**	**3**	**12**	**0.011**
ENSGACG00000017287	Chr3	13454527	13465167	*mmp16b*	12	0	0	0	12	0	0.001
ENSGACG00000017584	Chr3	14690814	14694448	*CCNY*	12	12	12	0	0	0	0.001
ENSGACG00000018249	Chr4	12141625	12143011	*si:ch211-153b23.5*	12	1	1	0	3	8	0.188
ENSGACG00000008034	Chr6	9368187	9380941		11	10	10	0	0	1	0.014
ENSGACG00000009469	Chr1	9166576	9173856	*egln2*	11	0	0	0	11	0	0.001
ENSGACG00000004433	Chr17	2127457	2211376	*igsf21a*	10	10	10	0	0	0	0.003
ENSGACG00000007343	Chr10	10666995	10679875	*col9a2*	10	0	0	0	6	4	0.227
ENSGACG00000018407	Chr4	13828336	13837518	*Sncb*	10	2	2	0	5	3	0.848
Salinity increase:											
ENSGACG00000020323	Chr7	17010160	17011176		23	0	0	0	22	1	<0.001
ENSGACG00000013229	**Chr18**	**15327717**	**15352321**		**15**	**10**	**10**	**0**	**1**	**4**	**0.125**
ENSGACG00000013359	Chr11	12960883	12968110	*sec14l1*	15	0	0	0	12	3	0.011
ENSGACG00000019416	**Chr7**	**4451892**	**4453656**	***HMX1 ortholog***	**15**	**3**	**3**	**0**	**5**	**7**	**0.745**
ENSGACG00000002948	Chr8	218240	221355	*ddx10*	14	0	0	0	6	8	0.077
ENSGACG00000016350	Chr14	3603545	3604923		14	1	0	1	7	6	0.277
ENSGACG00000006636	Chr18	4780893	4786820	*ZC3H12D*	13	0	0	0	3	10	0.034
ENSGACG00000004667	Chr12	4273498	4286193	*tti1*	12	0	0	0	12	0	0.001
ENSGACG00000015566	Chr2	9043062	9051779	*casc4*	10	0	0	0	10	0	0.003

respond to salinity change, which is expected for populations from different salinities.

Stable and inducible pop-DMS are associated with different functional gene categories

Gene functions associated with stable pop-DMS (452 and 329 under salinity decrease and increase, respectively) were enriched not only for a number of fundamental biological processes such as DNA repair and chromosome separation (fig. S2) but also for osmoregulatory functions (e.g., ion channel activity; Fig. 3). Furthermore, under increased salinity, many metabolic processes were found among the stable pop-DMS (Fig. 3). Inducible pop-DMS were associated with genes (100 and 82 under salinity decrease and increase, respectively) that were primarily enriched for other osmoregulatory functions regulating, for example, ion transmembrane transport (Fig. 3 and fig. S2). Therefore, stable and inducible pop-DMS affect not only different genes but also different gene ontologies with little overlap (Fig. 3 and fig. S2).

Assessing the role of inducible DNA methylation in nature

We investigated whether multiple pop-DMS associated with the same gene showed a correlated response to experimental salinity acclimation, which would require that they are nonrandomly distributed among the three categories stable, "inducible," and "inconclusive." Accordingly, we found a correlated response for pop-DMS at 13 of 20 genes (genes with more than 10 pop-DMS; Fisher's exact test, $P < 0.05$; Table 1), which suggests that inducible pop-DMS are predefined and directed.

We then tested whether inducible pop-DMS in the experimental fish became more similar to methylation levels found in natural

populations. Of the 198 (decreased salinity) and 148 (increased salinity) inducible pop-DMS, 130 (66%) and 101 (68%), respectively, became more similar to methylation levels of wild population to the corresponding salinity (hereafter referred to as "expected" direction). Conversely, at 68 (34%; decreased salinity) and 47 (32%; increased salinity) inducible pop-DMS, experimental fish showed methylation changes in the opposite direction, reducing the similarity to methylation levels observed in the natural populations (hereafter referred to as "opposite" direction).

Why, in a proportion of inducible methylation marks, the similarity between experimental and natural methylation levels was reduced was puzzling. One explanation could be a high level of genomic differentiation between the populations at these sites since genomic variation can have a strong cis-regulatory impact on epigenomic variation and may alter direction and function of methylation marks together (*21*). Thus, we hypothesized that opposite inducible pop-DMS are more often occurring in regions with higher genomic (DNA sequence–based) differentiation, while we anticipated the reverse at expected inducible pop-DMS. Accordingly, we resequenced whole genomes of the same wild-caught individuals that we used for RRBS and calculated the degree of genomic differentiation per inducible pop-DMS as mean F_{ST} value (±5-kb window) between populations. In line with our hypothesis, the populations from KIE (20 PSU) and NYN (6 PSU) were genetically more differentiated at opposite inducible pop-DMS than at expected sites (decreased salinity: δ.mean.F_{ST} = −0.014, P = 0.002; Fig. 4, A and C). A similar, yet not significant, trend was found between the populations from KIE (20 PSU) and SYL (33 PSU) (increased salinity: δ.mean.F_{ST} = −0.005, P = 0.153; Fig. 4, B and D). An alternative explanation is that not only salinity but also, rather, a combination of environmental cues (i.e., temperature, predation, and food) resulted in the methylation patterns found in the SYL population, which we did not include in our experiment. To understand whether selection has shaped the differences between increased and decreased salinity exposure, we tracked survival rates from fertilized eggs to the 3-month-old offspring and compared them between treatment groups. Mortality differed significantly between the treatment groups [generalized linear mixed model (GLMM), X^2_4 = 66.159, P < 0.001; Fig. 5A and table S3A] with increased mortality under increased salinity, while mortality under decreased salinity was generally low and did not differ from

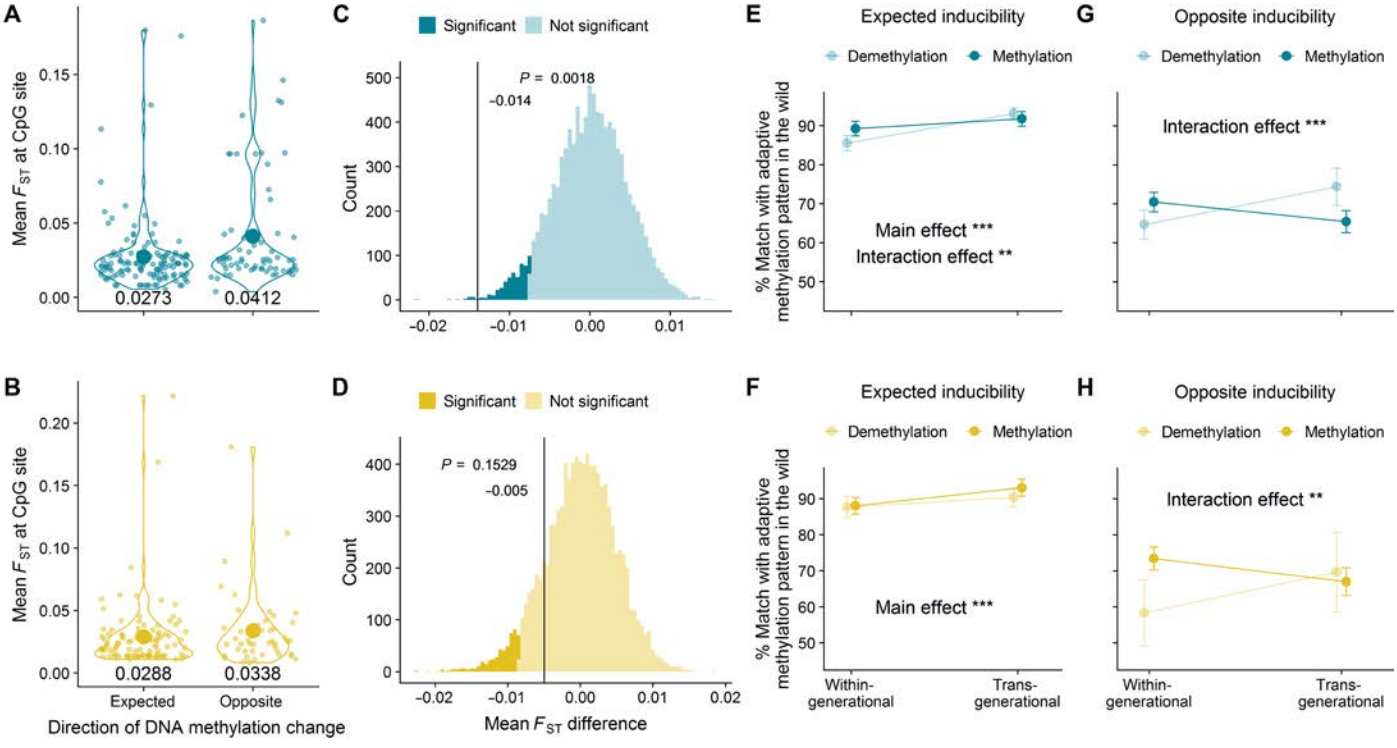


Fig. 4. The duration of acclimation (within-generational versus transgenerational) and level of genomic differentiation between populations influence DNA methylation at inducible sites. (**A** and **B**) Mean F_{ST} values for inducible pop-DMS (with a ± 5-kb window) under experimental salinity decrease (top; blue) and increase (bottom; yellow) that shifted methylation levels toward the values observed in either the field (expected) or the opposite direction (opposite). A randomization test (with 10,000 bootstraps) was performed for the difference between expected and opposite mean F_{ST} value (δ.mean.F_{ST} = expected mean F_{ST} − opposite mean F_{ST}) (**C** and **D**). Under the one-tailed hypothesis of increased genetic differentiation at opposite sites and an α of 0.05, the P value was calculated as values smaller than the true difference divided by 10,000 bootstraps. In (**E** to **H**), the y axis shows the percentage match between the within- and transgenerational acclimation groups in relation to the methylation differentiation level found in natural populations at inducible pop-DMS. This value was obtained by calculating the difference between the methylation change in the experiment (meth.diff.exp in %; control versus within-generational or control versus transgenerational) and the difference in methylation between natural populations (meth.diff.wild in %) as δ.meth.diff = 100 − (meth.diff.wild − meth.diff.exp). Mean values ± 95% confidence interval are shown for within- and transgenerational acclimation to decreased and increased salinity at expected and opposite inducible sites. Colors refer to the direction of DNA methylation change (hypomethylation or hypermethylation). Values closer to 100 indicate a shift in methylation pattern toward adaptive methylation levels found in natural populations, and asterisks indicate the significance level (***$P \leq 0.001$ and **$P \leq 0.01$) for the comparison between within- and transgenerational acclimation. "Main effect" refers to an effect of acclimation (within- or transgenerational), and "interaction effect" refers to an interaction of acclimation and methylation direction (hypo- or hypermethylation).

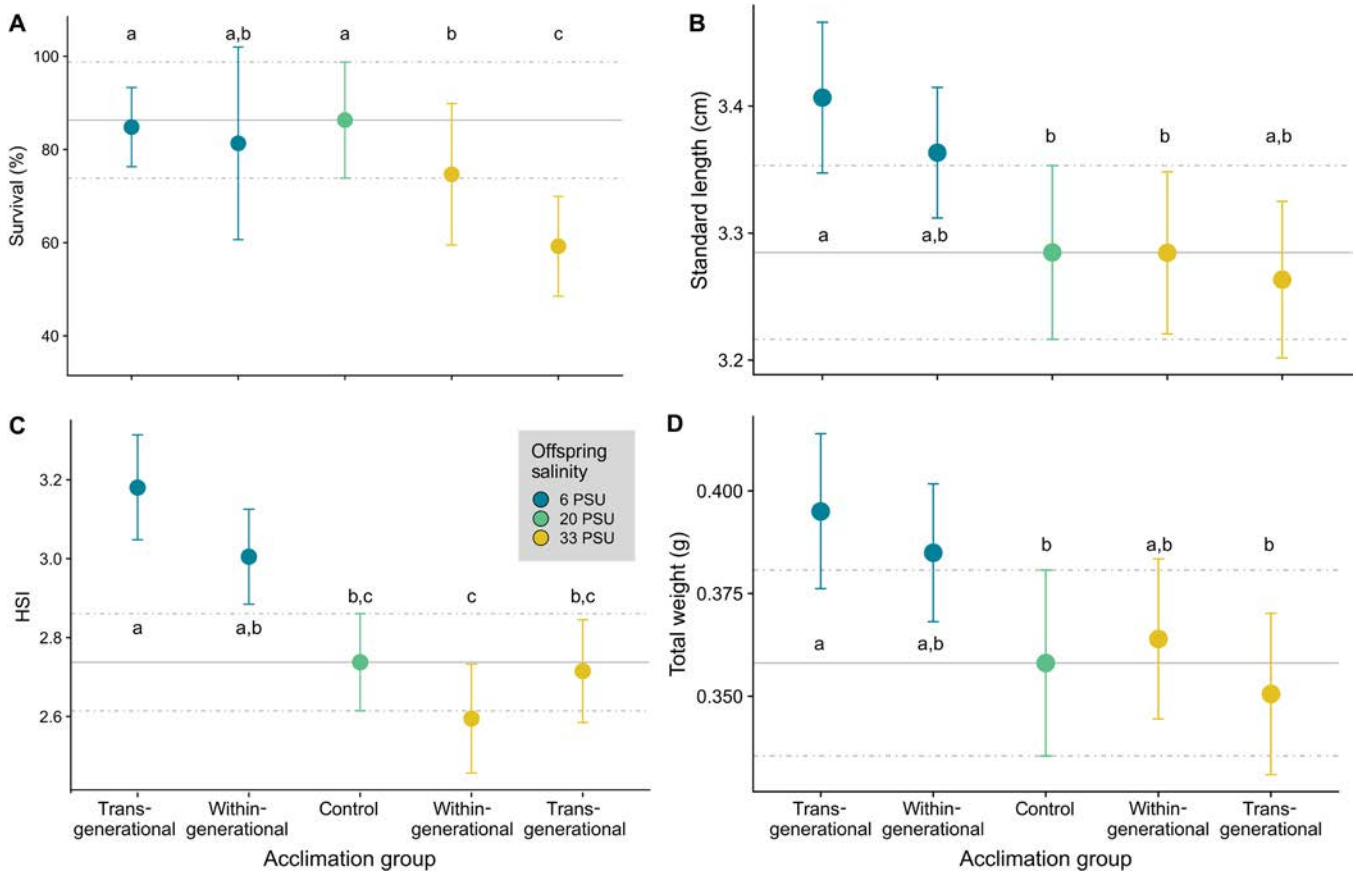


Fig. 5. Effects of salinity acclimation on fitness-correlated factors. For all five acclimation groups [control group (20 PSU), within-generational, and transgenerational acclimation to 6 or 33 PSU], survival rates in percent (**A**), standard length in centimeters (**B**), hepatosomatic index (**C**), and total weight in grams (**D**) are displayed. Letters indicate significant differences resulting from Tukey post hoc tests (table S3). HSI, hepatosomatic index.

the control group (Fig 5A and table S3A). Hence, while we cannot entirely disregard the effect of selection for increased experimental salinity, the patterns observed at pop-DMS upon reduced salinity are likely the sole result of tolerance mechanisms for salinity change.

Comparing within- and transgenerational acclimation effects on inducible DNA methylation

To test for adaptive transgenerational plasticity, we evaluated whether salinity acclimation over two, instead of only one, consecutive generations enhances the similarity of inducible pop-DMS with patterns found among wild populations at corresponding salinities. To do so, we calculated the percentage match (δ.meth.diff; Fig. 4, E to H) between the experimental groups and the anticipated adaptive methylation levels of wild populations. In line with our hypothesis, we found that transgenerational compared to within-generational salinity manipulation increased the δ.meth.diff (for expected inducible methylation, decreased salinity: $F_{1,256} = 30.42$, $P < 0.001$; increased salinity: $F_{1,198} = 10.39$, $P = 0.001$; Fig. 4, E and F). Under decreased experimental salinity, we found an interaction of "methylation direction" (hyper- or hypomethylation) and "acclimation" (within- and transgenerational) affecting the δ.meth.diff [analysis of variance (ANOVA), δ.meth.diff ~ methylation direction × acclimation, $F_{1,256} = 7.69$, $P = 0.006$; Fig. 4E]. Specifically, transgenerational acclimation increased the similarity of hypomethylated sites to methylation levels found in natural populations, while hypermethylated sites showed similar values within and across generations (Fig. 4E). While, for expected inducible sites, this effect was only present under decreased salinity, at opposite inducible sites, transgenerational acclimation to decreased and increased salinity elevated the δ.meth.diff at hypomethylated sites (ANOVA, δ.meth.diff ~ methylation direction × acclimation, decreased salinity: $F_{1,132} = 19.89$, $P < 0.001$; increased salinity: $F_{1,90} = 9.85$, $P = 0.002$; Fig. 4, G and H).

To infer the effect of DNA methylation differences on offspring, we compared fitness proxies among control, within-generational, and transgenerational acclimation groups (*22*). Specifically, we assessed the total weight, standard length (SDL), and the hepatosomatic index (HSI) as a proxy for energy reserves in the form of liver glycogen storage. SDL (GLMM, $X^2_4 = 9.965$, $P = 0.041$; Fig. 5B and table S3B) and total weight (GLMM, $X^2_4 = 11.518$, $P = 0.021$; Fig. 5D and table S3D) differed between treatment groups. Highly significant differences were detected for the HSI (GLMM, $X^2_4 = 22.688$, $P < 0.001$; Fig. 5C and table S3C), with elevated HSI observed under decreased salinity compared to fish from the control group. This supports previous findings, showing that osmoregulation at 6 PSU is energetically less demanding than that at higher salinities (*15*, *16*). Under increased salinity, HSI was lower compared to that of fish exposed to decreased salinity in the within-generational acclimation group, while a transgenerational acclimation to increased

salinity partially removed this difference. Although not significant, we observed a trend toward higher mean HSI in the transgenerational acclimation group compared to the within-generational acclimation group at the same salinity (Fig. 5C and table S3C).

DISCUSSION

This study investigated whether two postulated channels of epigenetic inheritance (selection based and detection based) can be identified in natural populations, focusing on salinity adaptation among populations of three-spined sticklebacks. Consistent with expectations for selection-based DNA methylation sites (*7*), we identified pop-DMS between populations that were both enriched for osmoregulatory functions and stable with respect to two generations of experimental salinity manipulation. Phenotypic variation originating from selection-based DNA methylation sites that are expected to have high epimutation rates [~10^{-4} for *Arabidopsis thaliana* (*9*)] could allow populations to explore the fitness landscape faster than under DNA sequence–based genetic variation alone [mutation rate, ~10^{-8} (*5*, *23*)]. Whether and at which rate these randomly emerging epimutations, as predicted for selection-based DNA methylation (*5*), occur in vertebrates remains unresolved. Notwithstanding, the observed enrichment of osmoregulatory gene functions for stable methylation sites (Fig. 3) suggests that they were subject to divergent natural selection, possible in interaction with DNA sequence–based variation. Furthermore, since local adaptation is 10 times more likely to involve changes in gene expression than in amino acid sequence (*24*), it is conceivable that differential DNA methylation and, consequently, regulation of osmoregulatory genes may contribute to local salinity adaptation. In sticklebacks, for instance, immunological adaptation has been shown to be mediated by gene expression (*25*). One of the top candidate genes differentially methylated between populations from 20 and 6 PSU was *eda* (ectodysplasin A), a well-described gene involved in lateral plate formation (*26*). Salinity and calcium are significant drivers of plate morphology (*27*) in proposed conjunction with predation (*28*). Our findings suggest that repeated and parallel selection for the low plated *eda* allele in response to low saline habitats (*29*–*31*), including the Baltic Sea (*14*, *32*), may also involve methylation-related mechanisms. Previous studies have shown that energetic cost for Baltic sticklebacks increases with increasing difference between treatment and isosmotic salinity conditions [~11 PSU (*33*)] (*15*, *16*). In line with these findings, we observed many metabolic processes associated with stable pop-DMS under increased salinity, also reflected in the lower HSI of fish at that salinity. Together, our results on the noninducible fraction of differentially methylated genes are consistent with a role in local salinity adaptation across stickleback populations (Fig. 3; Table 1; fig. S2; and table S2, A and B). These patterns of local adaptation in DNA methylation can have a genomic basis in the form of cis- and trans-acting genomic loci (*21*, *34*). Whether the differential methylation patterns represent an independent mechanism for local adaptation or are rather a consequence of DNA sequence–based genetic differentiation needs further study. Since our experiment only manipulated salinity while keeping all other factors constant, it is possible that some pop-DMS that were stable under salinity change could be inducible by other changing parameters.

With respect to the second postulated information channel, detection-based epigenetic inheritance (*7*), we identified more experimentally inducible pop-DMS than expected by chance. Multiple DMS associated with the same gene showed synchronized responses (Table 1). Furthermore, inducible pop-DMS were associated with different osmoregulatory genes compared to stable pop-DMS. Thus, inducible sites reflect a salinity-mediated plastic response, allowing individuals to regulate their ion balance relative to the seawater medium instantaneously without requiring any further genetic adaptation. More than two-thirds of these inducible pop-DMS became more similar to methylation patterns found in wild population. The similarity of these pop-DMS methylation levels between naturally adapted and experimentally acclimated population increased across generations. Considering the corresponding beneficial phenotypic effects, this strongly suggests that adaptive transgenerational plasticity plays a role in salinity acclimation. Since we used a split-clutch design for the breeding experiment, we can assume that these groups have similar genomic backgrounds. Furthermore, as mortality levels at low salinity remained low and did not differ between treatment groups, we can rule out any effect of selection altering the genotype composition in the groups at decreased salinity.

The induction of methylation sites has been discussed as a potential buffer for environmental changes (*11*, *17*, *35*). We found that the potential for adaptive transgenerational effects, specifically the ability to establish the anticipated adaptive methylation pattern found in the wild, differed among methylation directions (hypo- and hypermethylated sites; Fig. 4, E, G, and H), with a higher potential for transgenerational plasticity at hypomethylated sites. In line with our finding, the spontaneous addition of a methyl group to a cytosine is 2.5 times more likely than the removal (*23*). Methylation reprogramming that includes extensive methylation removal and de novo methylation during gamete formation and zygote development could thus serve as mechanisms to demethylate CpG sites in the transgenerational acclimation group (*36*, *37*).

The genetic background is considered to be an important source for epigenomic variation via cis- and trans-regulatory mechanisms (*21*, *38*, *39*). Thus, we characterized the genomic region surrounding each inducible pop-DMS and quantified the level of population differentiation (F_{ST}). This analysis revealed a negative correlation between population genetic differentiation and the propensity of the experimental population to approach the methylation level of the low salinity population (NYN) under salinity decrease (Fig. 4, A and C). Here, experimentally induced DNA methylation becomes more similar to the methylation in natural populations only in genomic regions with low genetic differentiation. On the other hand, when experimentally induced methylation differences to the low salinity population increase (Fig. 4, A and C), this occurs in a more divergent genomic background, suggesting that the genome has undergone selection leading to DNA-based local adaptation, rendering epigenetic modifications less relevant (*5*). Under increased salinity, a relationship between genomic differentiation (as F_{ST}) and methylation direction was inconclusive, suggesting that a combination of environmental cues shaped DNA methylation levels among wild populations at these sites. Overall, these findings emphasize the importance of the genomic background for interpreting DNA methylation patterns.

Together, our study provides the first empirical evidence that stable and inducible DNA methylation in wild animal populations follows predictions from evolutionary theory of selection- and detection-based epigenetic information channels (Fig. 2) (*5*, *7*). While the selection-based information channel assumes random variation from epimutation that is subsequently shaped by selection or drift, the detection-based information channel allows a directional response

in the form of transgenerational plasticity. Because the evolutionary implications of these two channels of inheritance are very different, future transgenerational or epigenetic studies should distinguish among both fundamentally different processes. Whether epigenetic marks, such as differentially methylated sites studied here, can permanently be attributed to one of the two categories or rather represent a continuum of stability levels and directionality will need further experimental testing over multiple generations.

MATERIALS AND METHODS

Animal welfare

All catches were performed under legal authorization issued by the German Ministry of Energy Transition, Agriculture, Environment, Nature and Digitalization in Schleswig-Holstein (MELUR: V242-7224.121-19), the Danish Ministry of Food, Agriculture and Fisheries of Denmark (case no: 14-7410-000227), the Estonian Ministry of the Environment (Keskkonnaministeerium - eripüügiluba nr 28/2014), and the Swedish Sea and Water Authority (Havs och Vattenmyndigheten). Ethical permission for the experiments required by German law was given by the MELUR: V312-7224.121-19, and the study is also in line with the Institutional Animal Care and Use Committee guidelines.

Survey and experimental design

For the field survey, we collected juvenile three-spined sticklebacks (*G. aculeatus*; 31.68 ± 14.25 mm) from three different salinity regimes inside and outside the Baltic Sea [SYL, Germany (55°00′58.3″N, 8°26′22.0″E), 33 PSU (n = 16); KIE, Germany (54°26′11.8″N, 10°10′20.2″E), 20 PSU (n = 16); NYN, Sweden (58°52′44.7″N, 17°56′06.2″E), 6 PSU (n = 16)] in September 2014. Fish were immediately euthanized using tricaine methane sulfonate solution (MS222), photographed, measured (length and total weight), and stored in RNAlater solution (24 hours at 7°C, afterward at −20°C). A cut along the ventral side ensured that the RNAlater solution diffused into all tissues. Conserved specimens were later dissected in the laboratory, and gill tissue was separated as the main osmoregulatory organ in fishes. For the acclimation experiment, we collected adult fish from KIE (20 PSU), which were crossed in our facilities at GEOMAR to obtain 10 F1 laboratory-bred families, herein referred to as "parental generation". At 9 months after hatch, we split each family into three salinity treatment groups of 10 fish each: one at 33 PSU, one at 6 PSU, and one control group at 20 PSU. The salinity transition was performed within 10 days by 3-PSU steps every second day. Over the entire time, each group was fed ad libitum and kept in a 20-liter aquarium connected to one of three filter tanks per salinity treatment. After 5 months under treatment conditions, six pure crosses per salinity treatment group were performed in vitro, herein referred to as "offspring generation" (F2). Offspring and parental generations were kept at 18°C water temperature and a 15:9 light/dark (L/D) cycle. During the past 8 weeks before the F2 crosses, the F1 generation underwent an artificial winter to trigger reproduction (2 weeks at 12°C, 12:12 L/D; 4 weeks at 6°C, 8:16 L/D; 2 weeks at 12°C, 12:12 L/D). Upon fertilization, clutches were split and separated into different treatments (Fig. 1). At 3 months after hatch, laboratory-bred F2 sticklebacks were euthanized using MS222, photographed, and dissected, and their gill tissue was stored in RNAlater solution. The age at sampling matched the estimated age of the wild-caught juveniles (3 months). In addition to the 48 wild-caught individuals from KIE, NYN, and SYL that were used in the above field survey, we sequenced whole genomes from gill tissue of an additional three populations of sticklebacks, namely, from Falsterbo, Sweden (55°24′46.6″N, 12°55′52.3″E; 10 PSU; n = 16), Letipea (59°33′07.6”N, 26°36′29.7″E; 4 PSU; n = 16), and Barsta (62°51′47.1″N, 18°23′51.0″E; 5 PSU; n = 16).

Mortality and HSI

Mortality was monitored throughout the experiment to account for possible nonrandom effect of selection. Three months after hatch, we assessed the SDL, total weight, and liver weight of the experimental F2 generation and calculated the HSI (HSI = liver weight/total weight × 100), which is a proxy for energy reserves in the form of glycogen storage. We analyzed the effect of treatment (five treatment groups; Fig. 1) on the survival rate per family as a ratio of "alive" versus "dead" fish using glmer implemented in the R package "lme4" (*40*) with binomial error and "crossing" as well as "climate chamber" as random effects. The effect of treatment on HSI, SDL, and total weight was analyzed fitting three individual linear mixed-effect models using lmer in lme4 (*40*) with Gaussian error and crossing as well as tank nested within climate chamber as random effects. Tukey post hoc tests were run using the glht function implemented in the package multicomp (*41*) to identify significant differences between treatment groups.

DNA extraction

For the field survey, DNA extraction of gill tissue (n = 16 individuals per population) was performed using the DNeasy Blood and Tissue Kit (QIAGEN). Further purification of the extracted DNA was done with NucleoSpin gDNA Clean-up (Macherey-Nagel). For laboratory-bred F2 offspring of the two-generation acclimation experiment, dual extraction of whole RNA and DNA was performed from gill tissue (n = 11 to 12 individuals per treatment group; Fig. 1) stored in RNAlater solution using the AllPrep DNA/RNA Mini Kit (QIAGEN). Purity and quality of the extracted DNA were estimated using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific) and a standard agarose gel (1% agarose/tris-acetate-EDTA). DNA concentration was assessed using the Qubit 2.0 Fluorometer (Thermo Fisher Scientific). To obtain a balanced sex ratio (50:50), we determined the gender of the individuals using a sex-specific genetic polymorphism in isocitrate dehydrogenase with a modified protocol from Peichel *et al.* (*42*). For the polymerase chain reaction (PCR) (settings: once 94°C for 3 min; 30 cycles of 94°C for 30 s, 54°C for 20 s, and 72°C for 30 s; once 72°C for 5 min), 1 μl of forward and reverse primer (5 μM) was used with 4.9 μl of water, 1 μl of 10× buffer, 1 μl of deoxynucleotide triphosphate (0.5 μM), and 0.1 μl of DreamTaq (5 U/μl). The resulting PCR products were visualized with a capillary electrophoresis on the 3100 ABI sequencer and a 500 LIZ size standard. While males show a heterogametic signal with two bands [at approximately 300 and 270 base pairs (bp)], females lack the band at 270 bp.

Library preparation and sequencing (whole-genome sequencing)

For whole-genome sequencing (WGS), the TruSeq Nano DNA (Illumina) library preparation kit was used according to the manufacturer's protocol by the Sequencing Facility of the IKMB, University of Kiel. Ultrasonication was conducted with a Covaris E220 (Covaris) to shear the input DNA (100 ng per sample and 350-bp insert size).

Before the enrichment with a PCR step (8 cycles), fragmented and bead-purified DNA was ligated with adenylate at the blunt 3′ ends (end repair and A-tailing) and with indexing adapters. Fragments were cleaned with MagSi-NGS^{prep} Plus Beads (Steinbrenner). Paired-end sequencing of the quality-controlled and multiplexed libraries was performed on the Illumina HiSeq 4000 platform (2 × 150–bp reads).

Quality assessment, data filtering, and mapping (WGS)

The command line tools of Picard version 2.7.1 (Broad Institute 2016) was used to (i) reformat the Fastq to uBAM file format and to add further values (read group, etc.) to the SAM header using FastqToSam, (ii) mark the location of adapter sequences using MarkIlluminaAdapters, and (iii) reconvert the sequences to Fastq format with SamToFastq. The stickleback genome (Broad/gasAcu1) was indexed with bwa index and used as a reference for the mapping with bwa mem (*43*) version 07.12-r1044. To retain the meta-information from the uBAMs, we used MergeBamAlignment. Picard was also used to identify duplicates with MarkDuplicates. Basic statistics were generated with CollectWgsMetrics, CollectInsertSizeMetrics, and AlignmentSummaryMetrics and summarized with MultiQC version 1.0.dev0 (*44*). A total number of 4,463,070,154 high-quality reads (mapping quality, >Q20) was mapped resulting in a mean depth of 13.84× (sd. 2.02×) and a mean insert size of 383.07 bp (sd. 9.40 bp; table S3). GATK version 3.7 HaplotypeCaller (*45*) was run to determine the likelihoods of the haplotypes per sample, i.e., to call single-nucleotide polymorphisms (SNPs) and indels (insertion-deletion), which were then processed with GenotypeGVCFs for a joint genotyping. SNPs were selected using hard filters for quality and extracted from the raw genotypes with a combination of the SelectVariants, VariantsToTable, and VariantFiltration commands. VCFtools (*46*) was used in a next step, removing SNPs with a minimum quality score below 20 and a minor allele frequency greater than or equal to 0.0049.

Library preparation and sequencing (RRBS)

The library preparation for methylation analyses followed the Smallwood and Kelsey RRBS protocol (*47*). A total of 100- to 250-ng purified DNA was digested with the methylation-insensitive Msp I restriction enzyme, which cuts at the "CCGG" motif and thereby enriches for CpG regions. DNA end-repair and A-tailing were conducted, and untailed CEGX spike-in controls (Cambridge Epigenetix) were added. These are DNA oligos of known sequence and with known cytosine modification, which can be used for downstream assessment of bisulfite conversion efficiency. After adapter ligation, bisulfite conversion was conducted using the EZ-96 DNA Methylation-Gold Kit (Zymo Research) according to the manufacturer's protocol. PCR amplifications with 19 cycles were performed. Quality control of purified PCR products was performed on a 2200 TapeStation System (Agilent), and high-quality libraries were pooled and diversified with 15% PhiX. Single-end sequencing with 100-bp read length was conducted on a HiSeq 2500 sequencer (Illumina).

Quality assessment, data filtering, and mapping (RRBS)

In total, 106 individuals (48 wild-caught and 58 experimental fish) of balanced sex ratio were DNA-sequenced at an average of 19.8 ± 3.5 million reads for experimental fish and 11.4 ± 2.1 million reads for wild-caught fish (table S4). Demultiplexed Fastq files were quality-checked using FastQC version 0.11.5 (*48*) and MultiQC version 1.3 (*44*). Adapters were removed with cutadapt version 1.9.1 (*49*) using multiple adapter sequences (NNAGATCGGAAGAG-CACAC, AGATCGGAAGAGCACAC, and ATCGGAAGAGCA-CAC) with a minimum overlap of 1 bp between adapter and read. This was necessary to remove primer dimers and avoid false methylation calls systematically caused by the RRBS end-repair step during library preparation, if the end-repair step adds artificial cytosines. Simultaneously, cutadapt was used to trim low-quality bases (-q 20) from the 3′ end and remove trimmed reads shorter than 10 bases. An air bubble during sequencing caused the bases 66 to 72 of 10 tiles of one lane (affecting 12 individuals) to have low-quality values, which were removed in a custom awk script. Two poor-quality individuals (a SYL and a NYN female) did not meet our strict quality requirements (e.g., ≥5 million reads; mapping efficiency, >52%) and showed biases in the proportion of bases per position compared to other individuals (plot in FastQC "per base sequence content"). Therefore, we excluded these two libraries from downstream analysis resulting in 15 instead of 16 individuals from SYL and NYN (Fig. 1). Bisulfite conversion efficiency was assessed from the spike-in controls (Cambridge Epigenetix) using the cegxQC software (*50*). Overall, conversion levels were 2.4 ± 1.8% conversion of methylated cytosines and 99.6 ± 0.5% conversion of unmethylated cytosines, which is in line with expected conversion rates (table S4). We used Bismark version 0.17.0 (*51*) to index the University of California Santa Cruz stickleback reference genome (Broad/gasAcu1) and to generate the bisulfite alignments with Bowtie2 version 2.3.3 at default settings. Bismark was also used to extract the methylation calls. Average mapping efficiency was 63.7 ± 2.4% (table S4).

Identification of differentially methylated sites

The methylation calls were analyzed in R version 3.4.1 (*52*) using the package methylKit version 1.3.8 (*53*). CpG loci were filtered for a minimum coverage of 10 reads per site. To account for potential PCR bias, we additionally excluded all sites in the 99.9th percentile of coverage. To improve the methylation estimates, we corrected for SNPs, which could have led to a wrong methylation call. The excluded positions were derived with custom-written Perl scripts from C-to-T and G-to-A SNPs with genotype quality of 20 and a minimum allele frequency of 0.005 (see above) from the 96 wild-caught individuals with a combination of custom-written Perl and R scripts using packages from methylKit (*53*) and GenomicRanges (*54*). After normalizing coverage values between samples, using normalizeCoverage implemented in methylKit, we excluded all sites that were present in fewer than nine individuals per treatment group from downstream analysis. As previously shown, sex-specific methylation affects <0.1% of CpG sites on autosomal chromosomes but >5% of CpGs on the sex chromosome (*18*). Therefore, to exclude a potential sex bias, we removed all CpG sites located on the sex chromosomes (chromosome 19), resulting in a high-quality dataset with 525,985 CpG sites. Last, by checking the first six principal components of the resulting principal components analysis and running an ANOVA on the filtered dataset, we confirmed the absence of an effect of sex on global methylation pattern ($F_{124,1} = 2.611$, $P = 0.109$). However, the principal components analysis revealed a bias in methylation pattern by families over all experimental groups. Therefore, to identify differentially methylated CpG sites (DMS) between treatment groups, we performed pairwise comparisons (table S5) fitting a logistic regression model per CpG site with calculateDiffMeth

in methylKit using family as covariate for the experimental groups. A chi-square test was applied to assess significance levels of DMS, and P values were corrected to q values for multiple testing using the sliding linear model method (*55*). In addition, we accounted for multiple use of groups in pairwise comparisons and adjusted the α for the q value according to Bonferroni correction to 0.0125 (0.05/4). Ultimately, CpG sites were considered to be differentially methylated with a $q < 0.0125$ and a minimum weighted mean methylation difference of 15%. To ensure that the DMS obtained are not laboratory artifacts, we used calculateDiffMeth implemented in methylKit and compared the wild population from KIE to the experimental control group (KIE population from 20 PSU at 20 PSU). The resulting 11,828 DMS were excluded from the DMS obtained by the pairwise comparisons mentioned above (table S5). DMS were plotted across the genome for the comparison between KIE versus NYN (20 versus 6 PSU; blue fish) and KIE versus SYL (20 versus 33 PSU; yellow fish) using ggplot2 (*56*) and hypoimg (*57*) (fig. S1).

Assessment of inducibility and gene association of DMS

By comparing wild-caught individuals from the mid-salinity population (20 PSU; KIE) to the populations sampled at low (6 PSU; NYN) and high (33 PSU; SYL) salinity in the field, we obtained 1470 (KIE-NYN) and 1158 (KIE-SYL) pairwise pop-DMS. We first tested whether these pop-DMS distinguishing natural populations are inducible or stable at the respective salinity in the experiment. A pop-DMS was considered stable when the within- and the transgenerational acclimation groups did not significantly differ in methylation to the control group ($q \geq 0.0125$). On the other hand, pop-DMS were considered inducible when at least one of the acclimation groups was differentially methylated compared to the control group ($q < 0.0125$; methylation difference, ≥15%). pop-DMS with a significant q value not exceeding the threshold of differential DNA methylation (15%) will be referred to as inconclusive hereafter. We used a randomization test to ensure that the number of inducible sites obtained did not occur by chance. To this end, we randomly sampled 1470 (KIE-NYN) and 1158 (KIE-SYL) pop-DMS from the complete dataset (1000 replicates). A chi-square test was used to assess whether our observed number of inducible, stable, and inconclusive sites differs from a random set of sites (averaged over replicates). Last, we tested whether the weighted mean methylation difference (meth.diff, in percentage) between wild populations matches the inducible methylation difference by subtracting the "meth.diff" in the experiment (exp) from the meth.diff between wild-caught populations (wild)

$$\delta.\text{meth.diff} = 100 - (\text{meth.diff.wild} - \text{meth.diff.exp})$$

As we subtracted this difference from 100, values closer to 100 indicated higher similarity of experimentally inducible methylation with the postulated adaptive DNA methylation pattern in natural populations. By comparing the "δ.meth.diff" for within- and transgenerational acclimation using an ANOVA, we can assess whether there is a difference in inducibility of methylation to match patterns found in wild-caught populations. All analyses were run separately for decreased (6 PSU; KIE-NYN) and increased (33 PSU; KIE-SYL) salinity.

To detect potential functional associations of the observed changes in DNA methylation state, we classified the genomic region of a pop-DMS on the basis of their nearest TSS using annotateWithGeneParts and getAssociationWithTSS implemented in genomation version 1.4.2 (*58*). We distinguished between promoter (1500 bp upstream and 500 bp downstream of TSS), exon, intron, and intergenic regions. To be associated to a gene, the pop-DMS had to be either inside the gene or, if intergenic, not further than 10 kb away from the TSS. We excluded three pop-DMS that were on a different reference scaffold and then the gene they were associated to on the chrUn linkage group (that merges scaffolds into one large artificial chromosome). Using the genes with associated pop-DMS, we applied a conditional hypergeometric Gene Ontology (GO) term enrichment analysis (false discovery rate–corrected $P \leq 0.05$) with the Ensembl stickleback annotation dataset "gaculeatus_gene_ensembl," and all genes that were associated to any sequenced CpG site were used as universe. We identified overrepresented biological processes, molecular functions, and cellular components using the packages GOstats version 2.5 (*59*) and GSEABase version 1.46 (*60*) and corrected for multiple testing using the false discovery rate method implemented in goEnrichment version 1.0 (*61*) in R version 3.6 (*52*). Figures were produced using ggplot2 version 3.2 (*56*).

Estimation of DNA sequence–based genetic differentiation at differentially methylated sites

To evaluate the genetic differentiation up- and downstream (in sum, 10 kb) of the pop-DMS position, we calculated the mean F_{ST} values (≤60% missing data and depth, ≥5) from WGS data of the exact same individuals with vcftools version 0.1.15 (*62*). We hypothesized that inducible CpG positions matching the methylation difference expected from the profile of the wild populations are genetically more similar between the populations than sites that changed in the opposite direction. To test this one-tailed hypothesis, we applied a randomization test (with 10,000 bootstraps) on the mean F_{ST} difference between the two groups (expected and opposite)

$$\delta.\text{mean}.F_{ST} = \text{expected mean } F_{ST} - \text{opposite mean } F_{ST}$$

We plotted the 10,000 delta mean F_{ST} values and calculated a P value by dividing the proportion of values smaller than the true difference by the number of bootstraps. Figures were produced using ggplot2 version 3.2 (*56*).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/12/eaaz1138/DC1

Fig. S1. Significant DMS throughout the genome for comparison between KIE versus NYN (20 versus 6 PSU; blue fish) and KIE versus SYL (20 versus 33 PSU; yellow fish).

Fig. S2. GO terms for biological processes, cellular components, and molecular functions under salinity increase (20 versus 33 PSU; yellow) and decrease (20 versus 6 PSU; blue) associated with pop-DMS.

Table S1. Relative distribution of DMS among genomic features.

Table S2A. Differentially methylated genes between populations from KIE (20 PSU) and NYN (6 PSU).

Table S2B. Differentially methylated genes between populations from KIE (20 PSU) and SYL (33 PSU).

Table S3A. Tukey post hoc test results for survival rate.

Table S3B. Tukey post hoc test results for SDL.

Table S3C. Tukey post hoc test results for HSI.

Table S3D. Tukey post hoc test results for total weight.

Table S4. Summary statistics for whole-genome resequencing of wild-caught sticklebacks.

Table S5A. Summary statistics for the RRBS of experimental fish.

Table S5B. Summary statistics for the RRBS of wild-caught fish.

Table S6. The number of DMS for each of the two pairwise population comparisons (pop-DMS).

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. E. Jablonka, The evolutionary implications of epigenetic inheritance. *Interface Focus* **7**, 20160135 (2017).
2. K. Laland, T. Uller, M. Feldman, K. Sterelny, G. B. Müller, A. Moczek, E. Jablonka, J. Odling-Smee, G. A. Wray, H. E. Hoekstra, D. J. Futuyma, R. E. Lenski, T. F. C. Mackay, D. Schluter, J. E. Strassmann, Does evolutionary theory need a rethink? *Nature* **514**, 161–164 (2014).
3. M. I. Lind, F. Spagopoulou, Evolutionary consequences of epigenetic inheritance. *Heredity* **121**, 205–209 (2018).
4. O. Bossdorf, C. L. Richards, M. Pigliucci, Epigenetics for ecologists. *Ecol. Lett.* **11**, 106–115 (2008).
5. F. D. Klironomos, J. Berg, S. Collins, How epigenetic mutations can affect genetic evolution: Model and mechanism. *Bioessays* **35**, 571–578 (2013).
6. P. A. Jones, Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**, 484–492 (2012).
7. N. Shea, I. Pen, T. Uller, Three epigenetic information channels and their different roles in evolution. *J. Evol. Biol.* **24**, 1178–1187 (2011).
8. I. Kronholm, S. Collins, Epigenetic mutations can both help and hinder adaptive evolution. *Mol. Ecol.* **25**, 1856–1868 (2016).
9. R. J. Schmitz, M. D. Schultz, M. G. Lewsey, R. C. O'Malley, M. A. Urich, O. Libiger, N. J. Schork, J. R. Ecker, Transgenerational epigenetic instability is a source of novel methylation variants. *Science* **334**, 369–373 (2011).
10. F. Johannes, E. Porcher, F. K. Teixeira, V. Saliba-Colombani, M. Simon, N. Agier, A. Bulski, J. Albuisson, F. Heredia, P. Audigier, D. Bouchez, C. Dillmann, P. Guerche, F. Hospital, V. Colot, Assessing the impact of transgenerational epigenetic variation on complex traits. *PLOS Genet.* **5**, e1000530 (2009).
11. T. Ryu, H. D. Veilleux, J. M. Donelson, P. L. Munday, T. Ravasi, The epigenetic landscape of transgenerational acclimation to ocean warming. *Nat. Clim. Change* **8**, 504–509 (2018).
12. D. H. Evans, Cell signaling and ion transport across the fish gill epithelium. *J. Exp. Zool.* **293**, 336–347 (2002).
13. T. B. H. Reusch, J. Dierking, H. C. Andersson, E. Bonsdorff, J. Carstensen, M. Casini, M. Czajkowski, B. Hasler, K. Hinsby, K. Hyytiäinen, K. Johannesson, S. Jomaa, V. Jormalainen, H. Kuosa, S. Kurland, L. Laikre, B. R. MacKenzie, P. Margonski, F. Melzner, D. Oesterwind, H. Ojaveer, J. C. Refsgaard, A. Sandström, G. Schwarz, K. Tonderski, M. Winder, M. Zandersen, The baltic sea as a time machine for the future coastal ocean. *Sci. Adv.* **4**, eaar8195 (2018).
14. B. Guo, J. DeFaveri, G. Sotelo, A. Nair, J. Merilä, Population genomic evidence for adaptive differentiation in baltic sea three-spined sticklebacks. *BMC Biol.* **13**, 19 (2015).
15. M. J. Heckwolf, B. S. Meyer, T. Döring, C. Eizaguirre, T. B. H. Reusch, Transgenerational plasticity and selection shape the adaptive potential of sticklebacks to salinity change. *Evol. Appl.* **11**, 1873–1885 (2018).
16. J. DeFaveri, J. Merilä, Local adaptation to salinity in the three-spined stickleback? *J. Evol. Biol.* **27**, 290–302 (2014).
17. L. N. S. Shama, F. C. Mark, A. Strobel, A. Lokmer, U. John, K. Mathias Wegner, Transgenerational effects persist down the maternal line in marine sticklebacks: Gene expression matches physiology in a warming ocean. *Evol. Appl.* **9**, 1096–1111 (2016).
18. D. C. H. Metzger, P. M. Schulte, The DNA methylation landscape of stickleback reveals patterns of sex chromosome evolution and effects of environmental salinity. *Genome Biol. Evol.* **10**, 775–785 (2018).
19. A. V. Artemov, N. S. Mugue, S. M. Rastorguev, S. Zhenilo, A. M. Mazur, S. V. Tsygankova, E. S. Boulygina, D. Kaplun, A. V. Nedoluzhko, Y. A. Medvedeva, E. B. Prokhortchouk, Genome-wide DNA methylation profiling reveals epigenetic adaptation of stickleback to marine and freshwater conditions. *Mol. Biol. Evol.* **34**, 2203–2213 (2017).
20. A. Bird, Perceptions of epigenetics. *Nature* **447**, 396–398 (2007).
21. A. Taudt, M. Colomé-Tatché, F. Johannes, Genetic sources of population epigenomic variation. *Nat. Rev. Genet.* **17**, 319–332 (2016).
22. F. Dufresne, G. J. FitzGerald, S. Lachance, Age and size-related differences in reproductive success and reproductive costs in threespine sticklebacks (*Gasterosteus aculeatus*). *Behav. Ecol.* **1**, 140–147 (1990).
23. A. van der Graaf, R. Wardenaar, D. A. Neumann, A. Taudt, R. G. Shaw, R. C. Jansen, R. J. Schmitz, M. Colomé-Tatché, F. Johannes, Rate, spectrum, and evolutionary dynamics of spontaneous epimutations. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6676–6681 (2015).
24. H. B. Fraser, Gene expression drives local adaptation in humans. *Genome Res.* **23**, 1089–1096 (2013).
25. T. L. Lenz, C. Eizaguirre, B. Rotter, M. Kalbe, M. Milinski, Exploring local immunological adaptation of two stickleback ecotypes by experimental infection and transcriptome-wide digital gene expression analysis. *Mol. Ecol.* **22**, 774–786 (2013).
26. P. F. Colosimo, K. E. Hosemann, S. Balabhadra, G. Villarreal Jr., M. Dickson, J. Grimwood, J. Schmutz, R. M. Myers, D. Schluter, D. M. Kingsley, Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* **307**, 1928–1933 (2005).
27. R. Spence, R. J. Wootton, M. Przybylski, G. Zięba, K. Macdonald, C. Smith, Calcium and salinity as selective factors in plate morph evolution of the three-spined stickleback (*Gasterosteus aculeatus*). *J. Evol. Biol.* **25**, 1965–1974 (2012).
28. A. Paccard, B. A. Wasserman, D. Hanson, L. Astorg, D. Durston, S. Kurland, T. M. Apgar, R. W. El-Sabaawi, E. P. Palkovacs, A. P. Hendry, R. D. H. Barrett, Adaptation in temporally variable environments: Stickleback armor in periodically breaching bar-built estuaries. *J. Evol. Biol.* **31**, 735–752 (2018).
29. A.-L. Ferchaud, S. H. Pedersen, D. Bekkevold, J. Jian, Y. Niu, M. M. Hansen, A low-density SNP array for analyzing differential selection in freshwater and marine populations of threespine stickleback (*Gasterosteus aculeatus*). *BMC Genomics* **15**, 867 (2014).
30. P. A. Hohenlohe, S. Bassham, P. D. Etter, N. Stiffler, E. A. Johnson, W. A. Cresko, Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLOS Genet.* **6**, e1000862 (2010).
31. N. V. Terekhanova, M. D. Logacheva, A. A. Penin, T. V. Neretina, A. E. Barmintseva, G. A. Bazykin, A. S. Kondrashov, N. S. Mugue, Fast evolution from precast bricks: Genomics of young freshwater populations of threespine stickleback *Gasterosteus aculeatus*. *PLOS Genet.* **10**, e1004696 (2014).
32. J. DeFaveri, P. R. Jonsson, J. Merilä, Heterogeneous genomic differentiation in marine threespine sticklebacks: Adaptation along an environmental gradient. *Evolution* **67**, 2530–2546 (2013).
33. T. Schaarschmidt, E. Meyer, K. Jürss, A comparison of transport-related gill enzyme activities and tissue-specific free amino acid concentrations of Baltic Sea (brackish water) and freshwater threespine sticklebacks, gasterosteus aculeatus, after salinity and temperature acclimation. *Mar. Biol.* **135**, 689–697 (1999).
34. M. J. Dubin, P. Zhang, D. Meng, M.-S. Remigereau, E. J. Osborne, F. Paolo Casale, P. Drewe, A. Kahles, G. Jean, B. Vilhjálmsson, J. Jagoda, S. Irez, V. Voronin, Q. Song, Q. Long, G. Rätsch, O. Stegle, R. M. Clark, M. Nordborg, DNA methylation in *Arabidopsis* has a genetic basis and shows evidence of local adaptation. *eLife* **4**, e05255 (2015).
35. G. Torda, J. M. Donelson, M. Aranda, D. J. Barshis, L. Bay, M. L. Berumen, D. G. Bourne, N. Cantin, S. Foret, M. Matz, D. J. Miller, A. Moya, H. M. Putnam, T. Ravasi, M. J. H. van Oppen, R. V. Thurber, J. Vidal-Dupiol, C. R. Voolstra, S.-A. Watson, E. Whitelaw, B. L. Willis, P. L. Munday, Rapid adaptive responses to climate change in corals. *Nat. Clim. Change* **7**, 627–636 (2017).
36. M. E. Potok, D. A. Nix, T. J. Parnell, B. R. Cairns, Reprogramming the maternal zebrafish genome after fertilization to match the paternal methylation pattern. *Cell* **153**, 759–772 (2013).
37. C. Labbé, V. Robles, M. P. Herraez, Epigenetics in fish gametes and early embryo. *Aquaculture* **472**, 93–106 (2017).
38. J. Gertz, K. E. Varley, T. E. Reddy, K. M. Bowling, F. Pauli, S. L. Parker, K. S. Kucera, H. F. Willard, R. M. Myers, Analysis of DNA methylation in a three-generation family reveals widespread genetic influence on epigenetic regulation. *PLOS Genet.* **7**, e1002228 (2011).
39. J. R. Gibbs, M. P. van der Brug, D. G. Hernandez, B. J. Traynor, M. A. Nalls, S.-L. Lai, S. Arepalli, A. Dillman, I. P. Rafferty, J. Troncoso, R. Johnson, H. R. Zielke, L. Ferrucci, D. L. Longo, M. R. Cookson, A. B. Singleton, Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLOS Genet.* **6**, e1000952 (2010).
40. D. Bates, M. Maechler, B. Bolker, S. Walker, lme4: Linear mixed-effects models using Eigen and S4. *R package version* **1**, 1–23 (2014).
41. T. Hothorn, F. Bretz, P. Westfall, Simultaneous inference in general parametric models. *Biom. J.* **50**, 346–363 (2008).
42. C. L. Peichel, J. A. Ross, C. K. Matson, M. Dickson, J. Grimwood, J. Schmutz, R. M. Myers, S. Mori, D. Schluter, D. M. Kingsley, The master sex-determination locus in threespine sticklebacks is on a nascent Y chromosome. *Curr. Biol.* **14**, 1416–1424 (2004).
43. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
44. P. Ewels, M. Magnusson, S. Lundin, M. Käller, MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
45. A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M. A. DePristo, The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
46. P. Danecek, A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin; 1000 Genomes Project Analysis Group, The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
47. S. A. Smallwood, G. Kelsey, Genome-wide analysis of DNA methylation in low cell numbers by reduced representation bisulfite sequencing, in *Genomic Imprinting: Methods and Protocols*, N. Engel, Ed. (Humana Press, 2012), pp. 187–197.
48. S. Andrews, FastQC: A quality control tool for high throughput sequence data (2010); www.bioinformatics.babraham.ac.uk/projects/fastqc.
49. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. Journal* **17**, 10–12 (2011).
50. CEGX Bioinfomatics Team, *Cambridge Epigenetix (CEGX)* (Babraham Research Campus, Cambridge, 2015).

51. F. Krueger, S. R. Andrews, Bismark: A flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572 (2011).
52. R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2017).
53. A. Akalin, M. Kormaksson, S. Li, F. E. Garrett-Bakelman, M. E. Figueroa, A. Melnick, C. E. Mason, methylKit: A comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* **13**, R87 (2012).
54. M. Lawrence, W. Huber, H. Pagès, P. Aboyoun, M. Carlson, R. Gentleman, M. T. Morgan, V. J. Carey, Software for computing and annotating genomic ranges. *PLOS Comput. Biol.* **9**, e1003118 (2013).
55. H.-Q. Wang, L. K. Tuominen, C.-J. Tsai, SLIM: A sliding linear model for estimating the proportion of true null hypotheses in datasets with dependence structures. *Bioinformatics* **27**, 225–231 (2011).
56. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
57. K. Hench, hypoimg (2019); https://github.com/k-hench/hypoimg.
58. A. Akalin, V. Franke, K. Vlahovíček, C. E. Mason, D. Schübeler, Genomation: A toolkit to summarize, annotate and visualize genomic intervals. *Bioinformatics* **31**, 1127–1129 (2014).
59. S. Falcon, R. Gentleman, Using GOstats to test gene lists for GO term association. *Bioinformatics* **23**, 257–258 (2006).
60. M. Morgan, S. Falcon, R. Gentleman, GSEABase: Gene set enrichment data structures and methods (2019).
61. A. Hallab, goEnrichment: Helper functions to compute GO enrichment tests using GOstats and GSEABase (2015).
62. B. S. Weir, C. C. Cockerham, Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370 (1984).
63. H. E. M. Meier, Baltic sea climate in the late twenty-first century: A dynamical downscaling approach using two global models and two emission scenarios. *Clim. Dynam.* **27**, 39–68 (2006).

Acknowledgments: We thank F. Wendt, J. Gismann, L. Sartoris, C. Giez, Z. Zagrodzka, L. Niewendick, M. Bettendorff, B. Poerschke, and T. Strickmann for taking care of the experimental animals and assisting the sampling procedures. We thank F. Rüppel for planning and constructing the recirculating aquaria systems. We thank the many field helpers and the BONUS BAMBI team for catching wild sticklebacks. We thank E. Trucchi for discussions about DNA methylation analysis. We also thank S. Landis who visualized our ideas in Figs. 1 and 2. **Funding:** We acknowledge the financial support for the BAMBI project (grant agreement number: call 2012-76) by BONUS; the joint Baltic Sea research program, funded by the European Union; and the Federal Ministry of Education and Research in Germany (to C.E. and T.B.H.R.) (reference number 03F0680A). Furthermore, B.S.M.'s research was supported by the Excellence Cluster "The Future Ocean" (EXC 80), and the sequencing was partly supported by the Excellence Cluster "Inflammation at Interfaces" (EXC 306) and "Origins and Function of Metaorganisms" (CRC1182), subprojects Z3 and INF, by the Deutsche Forschungsgemeinschaft (DFG). Both The Future Ocean and Inflammation at Interfaces are funded within the framework of the Excellence Initiative by the DFG on behalf of the German federal and state governments. **Author contributions:** C.E. and T.B.H.R. conceived and designed the study with contributions for the bisulfite sequencing strategy from B.S.M. B.S.M. and M.J.H. planned and carried out the fieldwork at the German locations. B.S.M. supervised the sampling in Estonia and Sweden, which was carried out by a great team of the BONUS-BAMBI project. M.J.H., with the help from C.E. and T.B.H.R., planned and supervised the breeding and acclimatization experiments. M.J.H. and B.S.M. conducted wet laboratory work (DNA extractions and quality assessment). M.P.H. and R.H. conducted the library preparation and sequencing. M.J.H. and B.S.M. analyzed the data and drafted the manuscript together with equal contributions. All coauthors discussed and interpreted the results and contributed to the final version of the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Fastq raw reads of RRBS were archived on NCBI [SRA accession: PRJNA590978 (experimental data) and PRJNA591803 (wild populations)]. Raw data on mortality and phenotypic effects as well as the blacklist and the vcf files produced from the WGS data are available at PANGAEA (phenotypes: https://doi.org/10.1594/PANGAEA.892493; genomic data: https://doi.org/10.1594/PANGAEA.909531). All other data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Scripts and additional data related to this paper may be requested from the authors.

Submitted 13 August 2019
Accepted 26 December 2019
Published 20 March 2020
10.1126/sciadv.aaz1138

Citation: M. J. Heckwolf, B. S. Meyer, R. Häsler, M. P. Höppner, C. Eizaguirre, T. B. H. Reusch, Two different epigenetic information channels in wild three-spined sticklebacks are involved in salinity adaptation. *Sci. Adv.* **6**, eaaz1138 (2020).

ScienceAdvances

Two different epigenetic information channels in wild three-spined sticklebacks are involved in salinity adaptation

Melanie J. Heckwolf, Britta S. Meyer, Robert Häsler, Marc P. Höppner, Christophe Eizaguirre and Thorsten B. H. Reusch

Sci Adv **6** (12), eaaz1138.
DOI: 10.1126/sciadv.aaz1138

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

McNew et al. BMC Evolutionary Biology (2017) 17:183
DOI 10.1186/s12862-017-1025-9

BMC Evolutionary Biology

RESEARCH ARTICLE Open Access

Epigenetic variation between urban and rural populations of Darwin's finches

Sabrina M. McNew[1], Daniel Beck[2], Ingrid Sadler-Riggleman[2], Sarah A. Knutie[1], Jennifer A. H. Koop[1], Dale H. Clayton[1] and Michael K. Skinner[2*]

Abstract

Background: The molecular basis of evolutionary change is assumed to be genetic variation. However, growing evidence suggests that epigenetic mechanisms, such as DNA methylation, may also be involved in rapid adaptation to new environments. An important first step in evaluating this hypothesis is to test for the presence of epigenetic variation between natural populations living under different environmental conditions.

Results: In the current study we explored variation between populations of Darwin's finches, which comprise one of the best-studied examples of adaptive radiation. We tested for morphological, genetic, and epigenetic differences between adjacent "urban" and "rural" populations of each of two species of ground finches, *Geospiza fortis* and *G. fuliginosa,* on Santa Cruz Island in the Galápagos. Using data collected from more than 1000 birds, we found significant morphological differences between populations of *G. fortis*, but not *G. fuliginosa*. We did not find large size copy number variation (CNV) genetic differences between populations of either species. However, other genetic variants were not investigated. In contrast, we did find dramatic epigenetic differences between the urban and rural populations of both species, based on DNA methylation analysis. We explored genomic features and gene associations of the differentially DNA methylated regions (DMR), as well as their possible functional significance.

Conclusions: In summary, our study documents local population epigenetic variation within each of two species of Darwin's finches.

Keywords: Epigenetics, *Geospiza*, Copy number variation, Galápagos Islands, DNA methylation

Background

Studies of the molecular basis of evolutionary change have focused almost exclusively on genetic mechanisms. However, recent work suggests that heritable modifications to gene expression and function, independent of changes to DNA sequence, may also be involved in the evolution of phenotypes [1–3]. One of the most common of these epigenetic mechanisms is DNA methylation, i.e. the chemical attachment of methyl groups (CH_3) to nucleotides (usually a cytosine followed by a guanine- "CpG") [4]. Methylation can be induced by the environment and affect gene expression and phenotypic traits without changing the DNA sequence itself [5–8]. Importantly, some patterns of methylation are heritable, meaning they have the potential to evolve [9–14]. Indeed, because DNA methylation modifications (epimutations) are more common than genetic mutations [15], they may play a role in the rapid adaptation of individuals to new or variable environments [16].

Environmentally-induced epimutations may be a component of the adaptive radiation of closely related species to new environments [17]. For example, Skinner et al. [18] showed that epigenetic variation is significantly correlated with phylogenetic distance among five closely related species of Darwin's finches in the Galápagos Islands. Although the adaptive significance of this epigenetic variation is unknown, some of the variants are associated with genes related to beak morphology, cell signaling, and melanogensis. The results of this study suggest that epigenetic changes accumulate over macroevolutionary time and further suggest that epigenetic changes may contribute to the evolution of adaptive phenotypes.

* Correspondence: skinner@wsu.edu
[2]Center for Reproductive Biology, School of Biological Sciences, Washington State University, Pullman, WA 99164-4236, USA
Full list of author information is available at the end of the article

Epigenetic variation also occurs among populations within single species [15, 19–24]. Some population epigenetic studies report correlations between methylation patterns and environmental factors, suggesting that differences in methylation are involved in local adaptation to different environments [21, 22, 24]. For example, in a study of populations of two salt marsh specialist plants living along a salinity gradient, Foust et al. [24] found that ground salinity is more closely correlated with epigenetic variation than genetic variation.

The purpose of our study was to investigate epigenetic variation between populations of each of two species of Darwin's finches: the medium ground finch *(Geospiza fortis)* and the small ground finch *(G. fuliginosa)* (Fig. 1). Darwin's finches are a closely related group of about 16 species endemic to the Galápagos Islands [25–28]. Long-term studies show rapid phenotypic changes in populations of finches in response to environmental pressures, including competition [26]. The molecular basis of these phenotypic changes is poorly known. Although recent genomic studies have identified alleles in several putative genes associated with beak size and shape [28–30], most genetic markers show little differentiation among populations or species [28, 30–34].

Epigenetic variation may contribute to the phenotypic diversity of Darwin's finch populations that cannot be detected through genomic studies. As an initial test of this hypothesis, we compared components of the morphology, genetics, and epigentics in populations of finches living at El Garrapatero, a relatively undisturbed locality, to populations living near Puerto Ayora (Academy Bay), the largest town in the Galápagos Islands (Fig. 1). Hereafter, we refer to these as the "rural" and "urban" sites, respectively. The two sites, which are only 10 km apart, are both arid, lowland scrub habitat along the south and south-eastern coast of the island. Vegetative cover, based on remote sensing spectroradiometric indicies, is slightly higher at the urban site; however, cover is highly correlated between the two sites year-round (Additional file 1: Figure S1). Despite the overall ecological similarity of the sites, anthropogenic disturbance at the urban site has increased dramatically over the past fifty years [35]. Observational studies suggest urbanization has effects on finch behavior and diet: birds in the urban population incorporate novel, human foods into their diets, whereas finches in the rural popuation do not [36]. To further explore potential impacts of urbanization of Puerto Ayora on ground finches, we tested for morphological, genetic, and

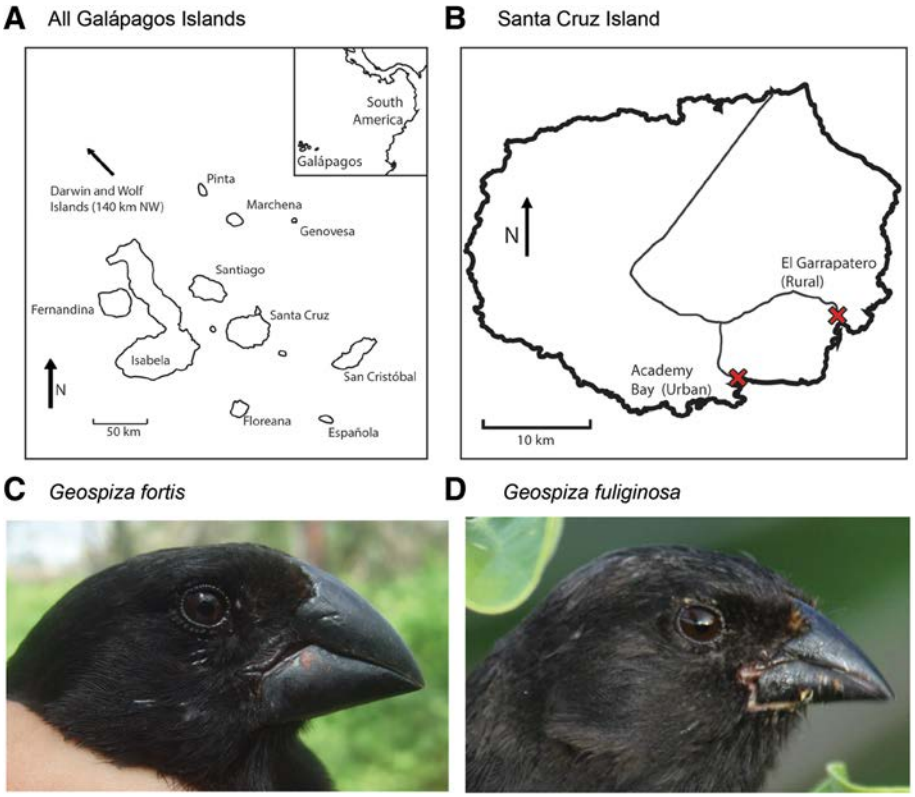


Fig. 1 Study sites and species. **a** The Galápagos Archipelago. **b** Santa Cruz Island; Roads are indicated by *narrow grey lines* and study sites by *red* Xs. **c** *Geospiza fortis*; photo by J.A.H.K. **d** *Geospiza fuliginosa*; photo by S.A.K. Maps in (**a**) and (**b**) are modified from © 2016 Google

epigenetic differences between urban and rural populations in each of two species of finches.

Methods

Study sites and species

We studied each of two populations of *G. fortis* and *G. fuliginosa* living in urban and rural environments (urban: Academy Bay; 0° 44′ 21.3″ S, 90° 18 ‘06.3″ W; rural: El Garrapatero; 0° 41′ 15.7″ S, 90° 13′ 18.3″ W). The two localities, which are separated by about 10 km, are both in the arid coastal zone of Santa Cruz Island (Fig. 1). *Geospiza fortis* and *G. fuliginosa* are among the most abundant species of finches at these study sites. There appears to be little movement of finches between populations. Over the course of a decade-long banding study (2002–2012), during which more than 3700 finches were captured- and more than 300 individuals recaptured- only one bird (a female *G. fortis*) was shown to have moved between the two sites (J. Raeymaekers pers. comm.).

Field work and sample collection

Finches were captured at the two study sites January–April 2008–2016. The birds were mist-netted and banded with individually numbered Monel bands in order to track individuals. They were aged and sexed using size and plumage characteristics [37]. Morphological measurements were taken from each individual including beak depth, beak width, beak length, tarsus length, wing chord, and body mass, following Grant and Grant (2014) [26], with the exception that wing chord was measured unflattened. Principle components were calculated from untransformed data for the three body measurements (mass, wing chord, and tarsus) and for the three beak measurements (length, width, and depth) to provide aggregate measures of body size and beak size and shape [38]. We evaluated morphological differences between urban and rural sites using linear mixed effects models (LMM), with site as a fixed effect, and year as a random effect to control for variation among years and investigators. Separate models were run for each morphological measurement, as well as body size (PC1 body) and beak size and shape (PC1 beak and PC2 beak). *P*-values were adjusted with a Bonferroni correction for multiple tests. Morphological analyses were run in the program RStudio using R version 3.2.1 with the packages pwr, plotrix, lme4, and lmerTest [39–42].

Blood and sperm samples for epigenetic and genetic analyses were collected from a subset of birds captured January–April 2009–2013 at the two study sites. Blood samples (<90 μl) were taken from finches via brachial venipuncture. The samples were stored on wet ice in the field and, within six hours of collection, erythrocytes were purified by centrifugation. Sperm samples (~5 μl) were taken from a subset of males. The sperm samples were obtained by gentle squeezing of the cloacal protuberance of reproductively active males. Blood erythrocytes and sperm samples were stored in a − 20 °C freezer in the Galápagos. Following each field season, they were transferred to a − 80 °C freezer in the USA for long-term storage. All field procedures were approved by the University of Utah Institutional Animal Care and Use Committee (protocols #07–08004, #10–07003 and #13–06010) and by the Galápagos National Park.

Genomic DNA preparation

Genomic DNA from finch red blood cells (erythrocytes) was prepared using the Qiagen DNAeasy Blood and Tissue Kit (Qiagen, Valencia CA). The manufacturer's instructions for nucleated blood samples were followed, but in the final DNA elution step H_2O was used instead of the buffer provided in the kit. Genomic DNA from finch sperm was prepared as follows: collected sperm suspension was adjusted to 100 μl with 1 x Phosphate Buffered Saline (PBS) then 820 μl DNA extraction buffer (50 mM Tris pH 8, 10 mM EDTA pH 8, 0.5% SDS) and 80 μl 0.1 M dithiothreitol (DTT) were added and the sample was incubated at 65 °C for 15 min. Next, 80 μl Proteinase K (20 mg/ml) were added and the sample was incubated on a rotator at 55 °C for 2 h. After incubation, 300 μl of protein precipitation solution (Promega, A795A) were added, then the sample was mixed and incubated on ice for 15 min, then spun at 4 °C at 13,000 rpm for 20 min. The supernatant was transferred to a fresh tube, then precipitated over night with the same volume of 100% isopropanol and 2 μl glycoblue at −20 °C. The sample was then centrifuged and the pellet washed with 75% ethanol, then air-dried and re-suspended in 100 μl H_2O. DNA concentration was measured using a Nanodrop Spectrophotometer (Thermo Fisher).

CNV-Seq protocol

To test for genetic differences between the urban and rural populations we sequenced DNA extracted from red blood cells (erythrocytes) and compared genetic copy number variation (CNV) [18]. CNV, defined as the changes in the number of repeat element copies of more than >1 kb of DNA, is increasingly recognized as one of the most common and functionally important markers of genetic variation [43]. The basic copy number variation (CNV) was determined through genomic sequencing of the same samples used for epigenetic analysis. Read numbers at specific loci were compared genome wide to identify CNV [18]. Erythrocyte DNA pools were generated by combining equal amounts of extracted DNA from five individuals. Each pool contained a total of 2 μg of genomic DNA. Three pools of five individuals each were created per species, per site.

Pooling samples for genomic analysis provides an accurate and cost-effective way of comparing populations [44]. Pooling decreases power, compared to sequencing individual samples. Although minor differences in copy number between populations may be missed [45], large differences between groups should be detected.

The pools were diluted to 130 μl with 1 x TE buffer and sonicated in a Covaris M220 with the manufacturer's preset program to create fragments with a peak at 300 bp. Aliquots of the pools were run on a 1.5% agarose gel to confirm fragmentation. The NEBNext DNA Library Kit for Illumina was used to create libraries for each pool, with each pool receiving a separate index primer. The libraries were sent to the University of Nevada, Reno Genomics Core for NGS on the Illumina HiSeq 2500 using a paired end PE50 application. All 6 pooled sequencing libraries for each species were run in one sequencing lane to generate approximately 30 million reads per pool. The read depth across the genome was then assessed to identify CNV and statistically assessed with a Bayesian analysis. The genome-wide paired end read depth was approximately 2× with the CNV read depth being a total of 300 to 6000 reads per CNV detected.

Methylated DNA Immunoprecipitation (MeDIP)

Following Skinner et al. [18], we used erythrocytes as a purified somatic cell type to compare differentially methylated regions (DMRs) between populations of each of the two species. For a subset of birds, we also compared DMR of germ line cells (sperm). DMRs between urban and rural populations were identified by the methylated DNA immunoprecipitation (MeDIP) of genomic DNA. MeDIP is an enrichment-based technique that uses an antibody to preferentially precipitate methylated regions of the genome that are then sequenced [46]. DMRs are identified by comparing coverage between groups of interest. MeDIP is a cost-effective way to evaluate genomic CpG methylation, and provides highly concordant results to other sequencing-based DNA methylation methods, such as bisulfite sequencing [47]. Because MeDIP surveys methylation genome-wide, it can be used to identify genomic characteristics associated with methylation. For instance, studies have found relationships between CpG density, methylation, and effects on gene transcription [6].

For analysis of erythrocytes, genomic DNA was extracted from the same individuals as used in the CNV pools. Each erythrocyte pool included five individuals and contained a total of 6 μg of genomic DNA. Sperm pools included two individuals and contained a total of 1.8 μg of genomic DNA. Three pools were generated per species, per site (for a total of $n = 6$ individuals per species, per site for sperm and $n = 15$ individuals per species per site for erythrocytes to consider biological variation of the pools and analysis). All pools were diluted to 150 μl with 1× Tris-EDTA (TE, 10 mM Tris, 1 mM EDTA) and sonicated with a probe sonicator using 5 × 20 pulses at 20% amplitude. Fragment size (200–800 bp) was verified on 1.5% agarose gel. Sonicated DNA was diluted to 400 μl with 1xTE and heated to 95 °C for 10 min, then shocked in ice water for 10 min. Next, 100 μl of 5 x immunoprecipitation (IP) buffer (50 mM Sodium Phosphate pH 7, 700 mM NaCl, 0.25% Triton X-100) and 5 μg of 5-mC monoclonal antibody (Diagenode, C15200006–500) were added and the sample was incubated on a rotator at 4 °C over night. The next day Protein A/G Agarose Beads from Santa Cruz Biotechnology, Santa Cruz CA, were prewashed with 1xPBS/0.1% BSA and re-suspended in 1 x IP buffer. Eighty μl of the bead slurry were added to each sample and incubated at 4 °C for 2 h on a rotator. The bead-DNA-antibody complex was washed 3 times with 1 x IP buffer by centrifuging at 6000 rpm for 2 min and resuspending in 1 x IP buffer. After the last wash the bead-complex was re-suspended in 250 μl of digestion buffer (50 mM Tris pH 8, 10 mM EDTA pH 8, 0.5% SDS) with 3.5 μl Proteinase K (20 mg/ml) per sample and incubated on a rotator at 55 °C for 2 h. After incubation, DNA was extracted with the same volume of Phenol-Chloroform-Isoamylcohol, then with the same volume of chloroform. To the supernatant from chloroform extraction, 2 μl glycoblue, 20 μl 5 M sodium chloride and 500 μl 100% cold ethanol were added. DNA was precipitated at −20 °C over night, then spun for 20 min at 13,000 rpm at 4 °C, washed with 75% ethanol, and air-dried. The dry pellet was re-suspended in 20 μl H_2O and concentration measured in Qubit using a Qubit ssDNA Assay Kit (Life Technologies, Carlsbad, CA).

MeDIP-Seq protocol

The next step for DMR identification involved sequencing the MeDIP DNA to identify differential methylation at specific genomic loci by assessing read numbers for the different samples. The MeDIP pools were used to create sequencing libraries for next generation sequencing (NGS) at the University of Nevada, Reno Genomics Core Laboratory using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina®, starting at step 1.4 of the manufacturer's protocol to generate double stranded DNA. After this step the manufacturer's protocol was followed. Each pool received a separate index primer. NGS was performed at the same laboratory using the Illumina HiSeq 2500 with a paired end PE50 application, with a read size of approximately 50 bp and approximately 100 million reads per pool. Two separate sequencing libraries, one rural and one urban, were run in each lane. The read depth for identified differential DNA methylated regions (DMRs) ranged from approximately 100 to >1000 total reads per DMR.

Bioinformatics

Basic read quality was verified using summaries produced by the FastQC program [48]. The reads for each sample for both CNV and DMR analyses were mapped to the zebra finch (*Taenopygia guttata*) genome using Bowtie2 [49] with default parameter options. The mapped read files were converted to sorted BAM files using SAMtools [50]. The cn.MOPS R package [51] was used to identify potential CNV. The cn.mops default information gain thresholds were used for this analysis. The cn.MOPS analysis detects CNVs by modeling read depth across all samples. The model predicts copy number for a given window based on observed read counts. The model uses a Bayesian framework to determine whether copy number for a give window differs significantly from 2. The length of the CNV is determined by comparing copy number of adjacent windows on the genome and joining those with the same copy number into one segment. A CNV call occurs when copy number for a given genomic segment varies from that of other samples. CNV detection with cn.MOPS is robust to low-coverage sequencing data (0.18–0.46 for 75 bp reads) and performs well when comparing 6 or more samples [51]. The window size used by the cn.MOPS analysis was chosen dynamically for each chromosome based on the read coverage. The chromosomes' window size ranged from approximately 25 kb to 60 kb. Only CNV that occurred in either all urban or all rural pools were compared. Although some individual pools had higher numbers of CNV, only CNV that occur red among all the pools were included in the analysis. The CNV are identified using the difference between the posterior and prior distributions from the Bayesian analysis to estimate information gain.

To identify DMR, the reference genome was broken into 100 bp windows. The MEDIPS R package [52] was used to calculate differential coverage between the urban and rural localities. The edgeR *p*-value [53] was used to determine the relative difference between the two localities for each genomic window. Windows with an edgeR *p*-value less than 10^{-3} were considered DMR. The DMR edges were extended until no genomic window with an edgeR *p*-value less than 0.1 remained within 1000 bp of the DMR. The DMR that included at least two windows with an edgeR *p*-value $<10^{-3}$ ("multiple-window DMR") were then selected for further analysis. Because no fully assembled or annotated genome exists for any Darwin's finch species, we aligned DMR with the zebra finch genome. CpG density and gene associations were then calculated for the DMR, based on alignment with the reference genome. Though we previously found high (>98%) homology between Darwin's finch and zebra finch genomes using tiling arrays [18], some differences were expected. Thus, associations of DMR with genes are likely to be under-estimates. To validate the epigenetics and gene associations, a similar analysis was also done with the draft *G. fortis* genome [54]. All the DMR sequence and genomic data obtained in the current study have been deposited in the NCBI public GEO database (GEO # GSE87825).

DMR clusters were identified with an in-house R script (www.skinner.wsu.edu under genomic data) using a 2 Mb sliding window with 50 kb intervals. DMR were annotated using the biomaRt R package [55] to access the Ensembl database [56]. The genes that overlapped with DMR were then input into the KEGG pathway search [57, 58] to identify associated pathways. A 10 kb flanking sequence was added to each DMR to consider potential localization in promoter regions of the gene as previously described [18, 59]. The DMR associated genes were manually sorted into gene classification groups by consulting information provided by the DAVID, Panther, and Uniprot databases incorporated into an internal curated database (www.skinner.wsu.edu under

Table 1 Mean (± 1SE) values for morphological characteristics of *G. fortis* and *G. fuliginosa* at rural vs. urban sites.

	G. fortis		*G. fuliginosa*	
Morphological	Rural	Urban	Rural	Urban
Character	*N* = 560	*N* = 245	*N* = 171	*N* = 121
Beak depth	11.48 ± 0.06	11.98 ± 0.09**	7.40 ± 0.04	7.42 ± 0.06
Beak width	9.89 ± 0.04	10.24 ± 0.07**	6.8 ± 0.03	6.82 ± 0.04
Beak length	11.71 ± 0.04	12.02 ± 0.07***	8.56 ± 0.04	8.46 ± 0.09
Tarsus length	21.00 ± 0.06	21.15 ± 0.09	18.83 ± 0.11	18.67 ± 0.09
Wing chord	69.3 ± 0.19	70.4 ± 0.29**	61.26 ± 0.31	61.1 ± 0.30
Body mass	21.23 ± 0.13	22.2 ± 0.23*	13.87 ± 0.15	13.76 ± 0.14
PC1 Body	−0.13 ± 0.06	0.29 ± 0.09***	0.07 ± 0.09	−0.10 ± 0.10
PC1 Beak	−0.17 ± 0.07	0.40 ± 0.11***	0.01 ± 0.09	−0.01 ± 0.15
PC2 Beak	−0.01 ± 0.02	0.02 ± 0.03	0.07 ± 0.04	−0.09 ± 0.10

Statistically significant differences between populations at $P < 0.01$, 0.001, and <0.0001 are indicated by *, ** and ***, respectively

Table 2 Differentially methylated regions (DMR) between urban and rural populations based on different cell types

Species/Cell Type	Number of Windows*					Sum of Multiple (≥2) Window DMR**
	1	2	3	4	5	
G. fortis Erythrocytes	2742	125	4	0	0	129
G. fortis Sperm	1160	97	9	3	1	110
G. fuliginosa Erythrocytes	4339	314	9	1	0	324
G. fuliginosa Sperm	1765	133	6	0	0	139

Only DMR that were significant at $P < 0.001$ are included
*DMR detected in one window alone were considered "single-window" variants (Fig. 2)
**DMR detected in two or more adjacent windows were considered "multiple-window" variants and used in subsequent analyses (Figs. 2, 3, 4, 5 and 6)

genomic data). To assess that the DMR were not false positives due to random biological variation within populations, a pairwise comparison analysis (individual pool comparison) on the genomic sequence data within the individual urban or rural sites and cell populations was performed [60].

Results

Morphology

We used 1097 birds captured between 2008 and 2016 for morphological analyses. We controlled for slight variation among years in traits by including year as a random effect in all analyses. At both sites, *G. fortis* was significantly larger than *G. fuliginosa* in all morphological traits (linear mixed-effects models $P < 0.0001$). Within species, urban *G. fortis* was significantly larger than rural *G. fortis* for all direct morphological measurements, except tarsus length (Table 1). Composite measures of *G. fortis* body and beak size (PC1 body and PC1 beak) also differed between the two sites; however, there was no difference in beak shape (PC2 beak). In contrast to *G. fortis*, *G. fuliginosa* did not differ significantly between the urban and rural populations in any of the morphological measurements or composite measures (Table 1). Because we captured more *G. fortis* than *G. fuliginosa* we did a power analysis for *G. fuliginosa*, using the effect size of morphological differences found in the *G. fortis* populations (smallest effect size = 0.256 (wing chord); largest effect size = 0.358 (beak depth)). Power for comparisons of *G. fuliginosa* appeared adequate for detecting similar effect sizes (0.69–0.91).

Copy number variation (CNV)

Mean read depth genome-wide for pools used in CNV analysis varied between 1.08× and 1.30× (overall mean = 1.22×). The total read depth of individual variants ranged from 300 to 6000. We identified unique CNV in three of six *G. fortis* pools and five of six *G. fuliginosa* pools. The total number of variants per pool ranged from 1 to 20. However, no variants were exclusive to all urban or all rural pools for either *G. fortis* or *G. fuliginosa* (Additional file 2: Fig. S2).

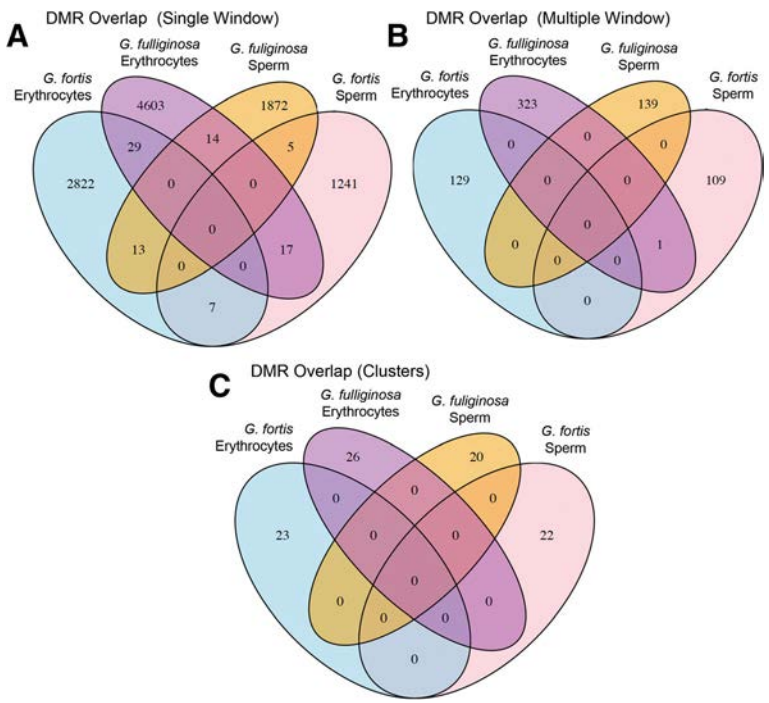


Fig. 2 DMR overlap between species and cell types. Each value is the number of differentially methylated regions between the urban and rural populations. Overlapping colors in the figure show the number of DMR that are shared between the two species or the two cell types. DMRs detected within **a** single 100 bp windows, **b** 2–5 adjacent 100 bp windows, **c** 2–7 Mb regions

Therefore, while there was variation within populations in copy number at various loci in both *G. fortis* and *G. fuliginosa* (e.g., FB2 & 12), there were no fixed differences between the urban and rural populations for either species. It is unclear why certain pools had more variants than others; however variation was consistent among chromosomes.

To control for underestimation of CNV differences due to reads that did not align to the zebra finch genome, we performed a similar analysis aligning reads to the un-assembled *Geospiza fortis* genome [54]. The average proportion of reads aligned to the *G. fortis* genome was higher (two-fold). However, we still did not find any differences in CNV between the urban and rural populations for either species of Darwin's finch. A limitation of this CNV analysis is that only large variants (>24 Kbp) can be detected reliably; smaller variants (<10 Kbp or less) may have escaped detection.

Differential DNA methylation regions (DMRs)

DMRs were found between populations for both cell types and both species (Table 2). We report the number of DMRs at *p*-value cut-offs ranging from <0.01 to <1e-05 in Additional file 3: Table S1; Additional file 4: Table S2 and Additional file 5: Table S3. The analyses reported below are restricted to DMRs significant at a level of $P < 0.001$. We evaluated differences on three "regional" scales (Fig. 2): single 100 bp window DMRs, multiple window DMRs, and "DMR clusters", i.e. statistically over-represented DMR clusters of 3–10 DMRs spanning 2–7 Mb [18] (Additional file 6: Table S4A-D). We focus on multiple-window DMRs (Additional file 4: Table S2 and Additional file 5: Table S3), i.e. DMRs detected independently in adjacent windows, because they further reduce the likelihood of false positives and provide a set of highly reproducible DMRs [18]. Multiple-window DMRs were used in the analysis of the genomic features of DMRs reported below.

There was little overlap between species or cell types in the regions that were differentially methylated between urban and rural populations (Fig. 2). A small proportion of single window DMRs (Fig. 2A) was shared between species and/or cell types. However, there were virtually no shared multiple-window DMRs (Fig. 2B) or clusters of DMRs (Fig. 2C) between species and/or cell types.

For both species and cell types, multiple-window DMRs usually were detected in only two multiple 100 bp windows; however, a limited number (<10% of total DMRs) were found in 3–5 multiple windows (Table 2). Based on extension of edges of multiple-window DMRs (extension

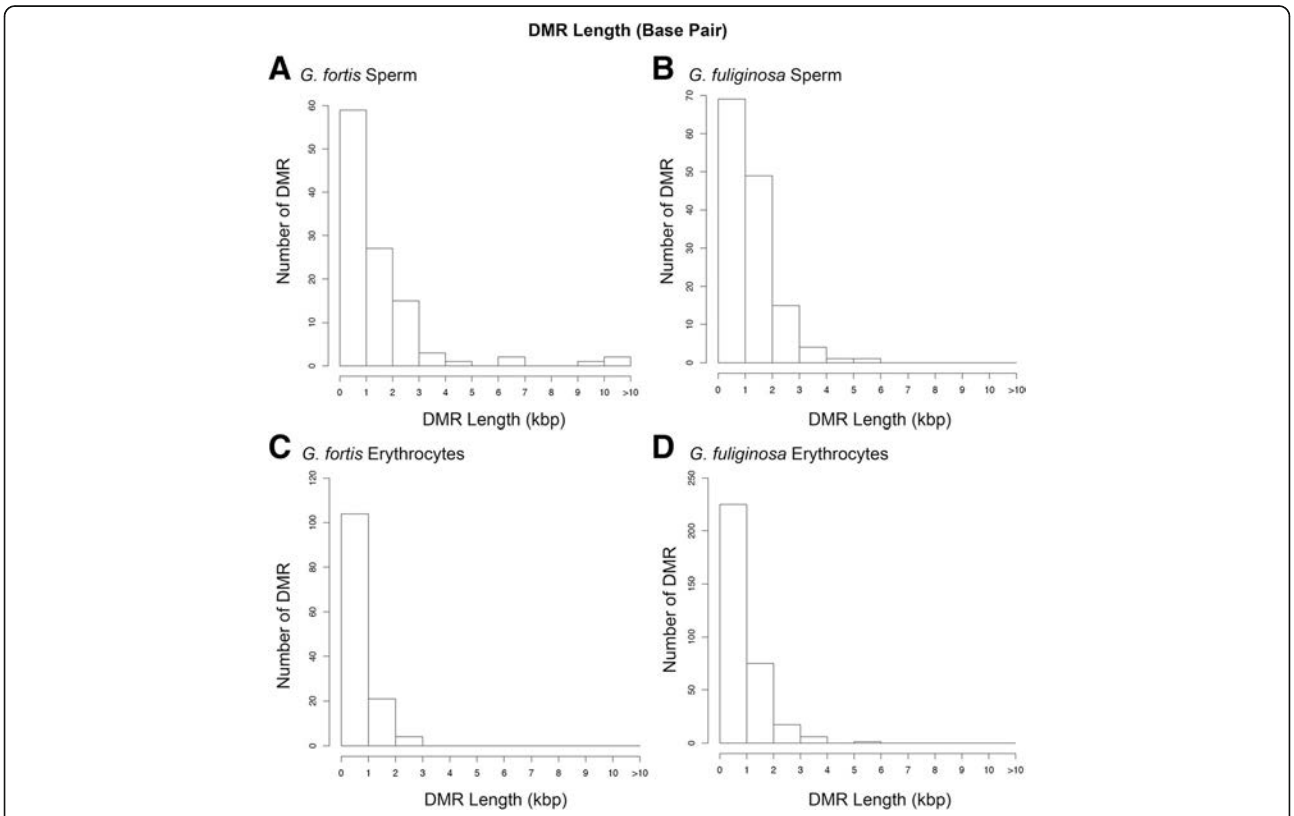


Fig. 3 DMR length (kb) in **a** *G. fortis* sperm. **b** *G. fuliginosa* sperm. **c** *G. fortis* erythrocytes. **d** *G. fuliginosa* erythrocytes. Only multiple-window DMR significant at a *p*-value threshold of $<10^{-3}$ are included

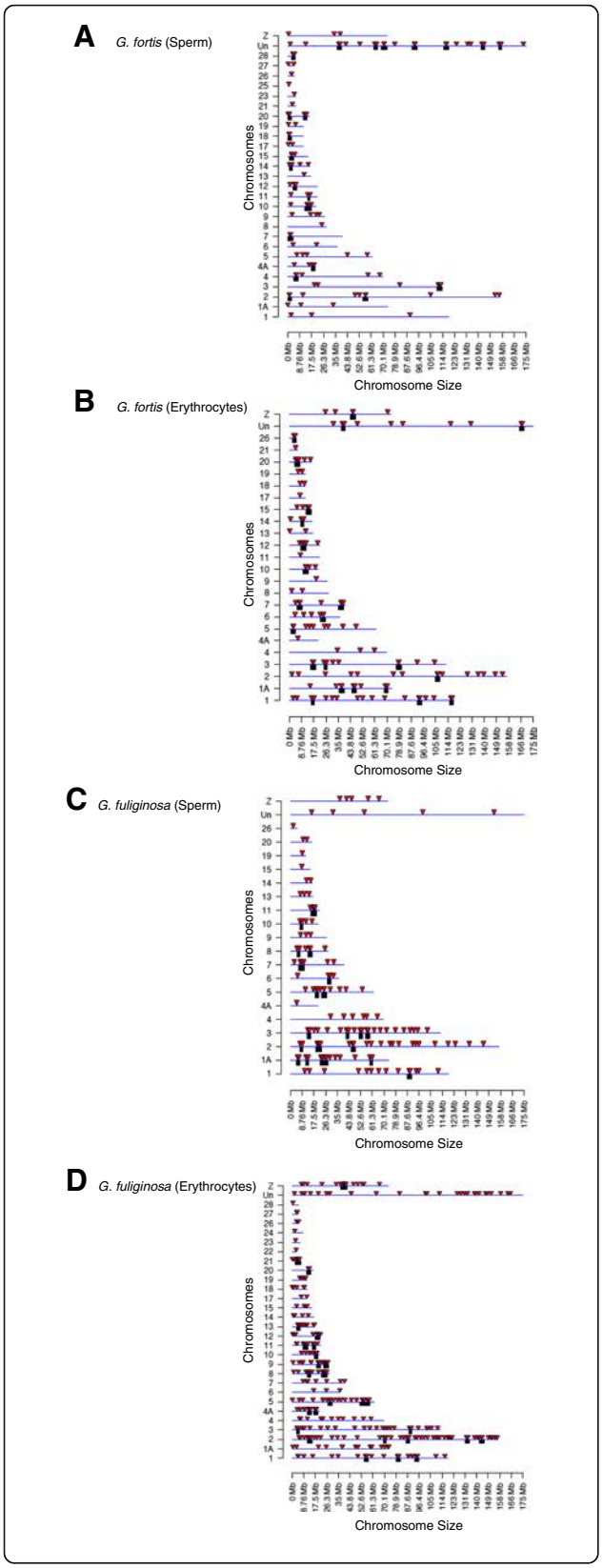


Fig. 4 Chromosomal locations of DMR identified in *Geospiza fortis* sperm **a** and erythrocytes (**b**) and *G. fuliginosa* sperm (**c**) and erythrocytes (**d**)**.** Locations are based on alignment to the zebra finch (*Taeniopygia guttata*) genome. *Red arrowheads* indicate DMR and *black boxes* indicate DMR clusters. Only multiple-window DMR significant at a *p*-value threshold of $<10^{-3}$ are shown

of adjacent 100 bp windows with $p < 0.1$; see Methods) we estimated that most DMRs were 500–1000 bp in length (Fig. 3). Many of the DMRs in this study were clustered together, consistent with previous studies showing that DMRs are not evenly distributed across the genome [59]. Based on alignment to the zebra finch genome, we plotted the chromosomal locations of multiple-window DMRs and DMR clusters (Fig. 4). DMRs were present on all chromosomes in both sperm and erythrocytes of both species; however, the chromosomal locations of DMRs differed between the cell types and species.

We evaluated the location of DMRs with respect to nucleotide composition. CpG density was highest in DMRs of *G. fortis* sperm cells (Fig. 5A). DMRs in *G. fortis* erythrocytes and both cell types of *G. fuliginosa* were most often found in lower density CpG regions of the genome (<1 CpG site/100 bp; Fig. 5B-D). We estimated that the DMRs typically had approximately 10 CpG sites clustered within 1 kb regions.

We identified potential genes associated with DMRs through alignment with the zebra finch reference genome. DMRs within 10 kb of a gene (such that the promoter is included) have the potential to influence the gene's expression and/or pathways associated with that gene [59]. Different categories of genes were methylated in the two cell types and species (Fig. 6, specific genes listed in Additional file 7: Table S5). The most common gene categories associated with DMRs were metabolism, cell signaling and transcription (Fig. 6). Gene categories associated with DMRs differed significantly between the two species (Chi-square test, $p = 0.039$) and marginally between the two cell types (Chi-square test; $p = 0.078$). Pathway analysis (KEGG) showed DMRs associated with several genes (GALNT14, SGMS1, ENO2, PLCH2) in metabolic pathways of *G. fortis* sperm. DMRs were associated with different genes (GCLC, PRIM2, ALD1A3, AK4, ACACA) in metabolic pathways of *G. fuliginosa* sperm. *Geospiza fortis* erythrocyte DMRs were associated with genes (CACNA1H, FGF8, MRAS, RAP1A) in the MAPK signaling pathway. *Geospiza fuliginosa* erythrocyte DMRs were not associated with any particular pathway.

When the DMR data sets for both species and cell types were compared, KEGG pathways with the most DMR-associated genes were metabolic pathways, and MAPK and TGFß/BMP signaling pathways. Metabolic pathways included glycolysis, in which genes involved with pyruvate and acetate production were associated with

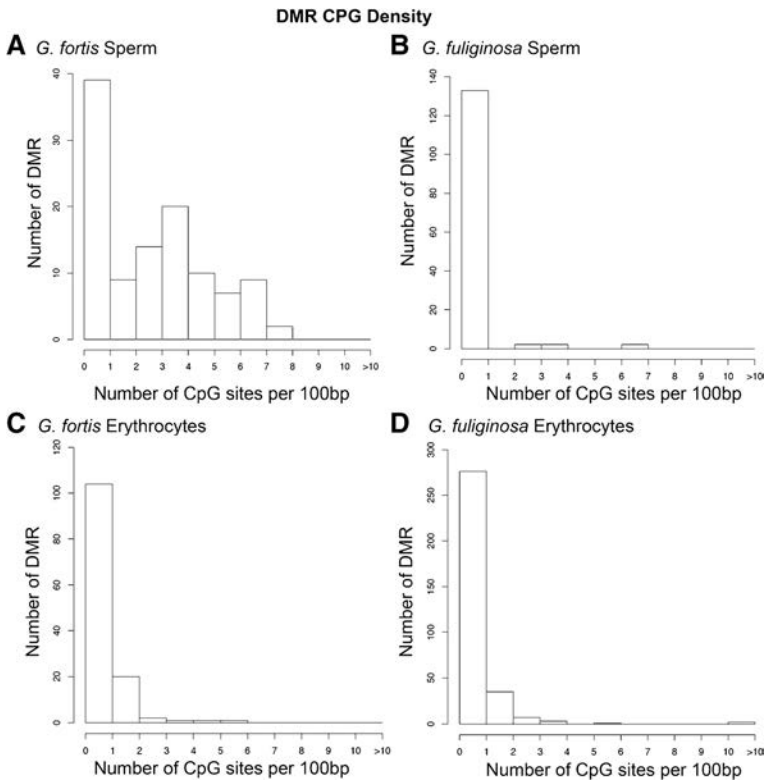


Fig. 5 The CpG density of DMR in *Geospiza fortis* sperm (**a**), *G. fuliginosa* sperm (**b**), *G. fortis* erythrocytes (**c**) and *G. fuliginosa* erythrocytes (**d**). Only multiple-window DMR significant at a *p*-value threshold of $<10^{-3}$ are included

DMRs in both finch species (Additional file 8: Figure S3 and Additional file 9: Figure S4). Other metabolic pathways associated with DMRs included genes involved in purine metabolism and glycosylation (Additional file 7: Tables S5). Signaling pathways were also associated with DMRs in both species and cell types. Three genes in the TGFß/BMP pathway were associated with DMRs between *G. fuliginosa* populations (erythrocytes and sperm combined): BMP5, BMP7 and FST (Fig. 7). MAPK, a common pathway for many regulatory processes, such as cell growth, contained a high number of DMR-associated genes in both finch species (Additional file 8: Figure S3 and Additional file 9: Figure S4).

Genomic correlates of our DMR and CNV data were analyzed using the well-annotated zebra finch genome. In addition, our sequencing data were also compared to the *G. fortis* shotgun sequence database [54]. In contrast to the zebra finch genome, the *G. fortis* genome is neither assembled, nor annotated, meaning that limited data analysis is possible. The pooled individual sample read number was approximately 100 million reads for both genome analyses. The overall read alignment rate was 47–48% for the zebra finch analysis and 70–75% for the *G. fortis* genome analysis. Although previous analysis using tiling arrays suggested a 98% similarity in tiling array hybridization of the genome [18], the next generation sequencing analysis shows that more differences exist, likely in non-coding regions. The zebra finch genome analysis revealed twice the number of DMRs compared to the *G. fortis* genome analysis. This was largely due to the incomplete nature of the *G. fortis* genome. Nevertheless, analysis with both the zebra finch and *G. fortis* genomes identified epigenetic alterations between the rural and urban sites. To test whether methylation variation between sites was greater than within sites we conducted a pairwise comparison analysis (comparison of individual pools) within each species and rural or urban populations for specific cell types. We identified a number of DMRs between individual pools, which suggests that there is epigenetic variation within the study populations. However, few DMRs were found in multiple pools from the same population. Moreover, almost none of these DMRs were also found between urban and rural populations (Additional file 10: Figure S5). Thus, the DMRs identified between urban and rural populations are not an artifact of sampling within-population variation.

Discussion

Darwin's finches are well known for their phenotypic variability and evolution in response to changing environmental conditions [26]. In addition to genetic variation,

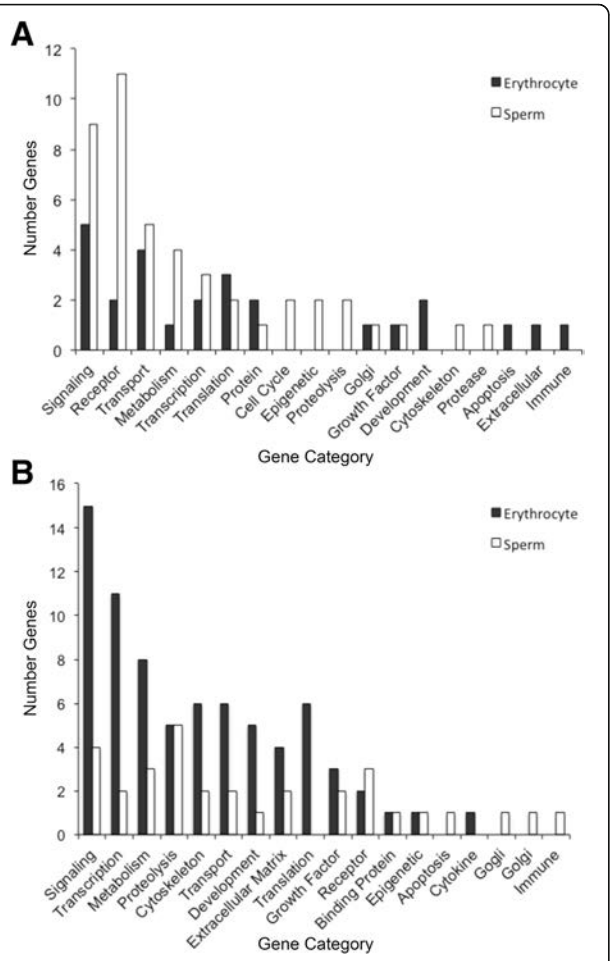


Fig. 6 Gene categories associated with DMR detected in (**a**) *G. fortis* and (**b**) *G. fuliginosa*. Only multiple-window DMR significant at a *p*-value threshold of $<10^{-3}$ are included

epigenetic variation - such as differential DNA methylation - may exist between natural populations living under different environmental conditions. The goal of this paper was to test for morphological, genetic, and epigenetic differences between urban and rural populations within each of two species of Darwin's finches. We found that *G. fortis* individuals at the urban site (Academy Bay) were larger than those at the rural site (El Garrapatero). In contrast, *G. fuliginosa* individuals did not differ morphologically between the sites. We did not find genetic differentiation between populations of either species based on CNV comparisons. However, we did find epigenetic (DMR) differences between urban and rural populations of both species of finches.

We found urban *G. fortis* were larger in nearly all morphological measurements compared to rural *G. fortis* (Table 1), which may be due to increased food availability at the urban site. Previous work suggests that urbanization around Academy Bay has relaxed selection on finch beak size [35, 36]. Urbanization is associated with a shift in the distribution of beak size in *G. fortis*: beak size is strongly bimodal at the rural site, whereas bimodality has decreased at the urban site concurrently with human population growth [35]. Both studies propose that increased food availability at the urban site has altered the selective landscape for *G. fortis* [35, 36]. Beak size is highly heritable in *Geospiza* finches; e.g. mid-parent vs. mid-offspring values estimate heritability of beak depth in *G. fortis* to be 0.74 [61].

In contrast, *G. fuliginosa* showed no morphological differentiation between sites (Table 1). *Geospiza fortis* is phenotypically more variable than *G. fuliginosa* on Santa Cruz Island [61]. As a result, *G. fortis* may have undergone more rapid local adaptation than *G. fuliginosa*. Although *G. fuliginosa* and *G. fortis* have overlapping dietary niches, they do show some degree of specialization [27]. It is possible that urbanization has had a greater selective effect on *G. fortis* than *G. fuliginosa*. Alternatively, morphological differences in *G. fortis* may be driven by hybridization between *G. fortis* and the slightly larger *G. magnirostris*. Hybridization between *G. fortis* and *G. magnirostris* has been documented on Santa Cruz [62]. While we have no information on relative rates of hybridization at our study sites, *G. magnirostris* is more abundant at the urban site than the rural site (4.56% of urban birds captured, compared to 1.86% of rural birds captured; unpublished data 2008–2016).

Despite differences in morphology between populations of *G. fortis*, we found no genetic differences between the urban and rural populations, based on the CNV comparisons made. Because CNV sequence coverage was limited, we may have overlooked small CNV, but larger CNV should have been detected between the two populations. CNV is a sensitive index of genetic differentiation between populations; indeed, some studies have found that CNV accounts for more genetic variation than SNPs [63–65]. Recent work has also linked CNV to rapid evolution in pepper moths [66] and primates [67].

Our study is first to explore genetic variation between populations of Darwin's finches using large–scale genomic features (CNV). Like our study, previous studies using smaller-scale genomic markers (microsatellites, nuclear introns, and mitochondrial DNA) detected little or no genetic structure within populations of either *G. fortis* or *G. fuliginosa* [31, 34, 68]. Two recent studies of genomic variation among Darwin's finches using SNPs did identify variable sites associated with variation in beak morphology [29, 30]. However, most of the genes associated with beak morphology in the two studies were different. These inconsistent results suggest that other forms of variation, such as large scale CNVs, could underlie phenotypic differences. However, our

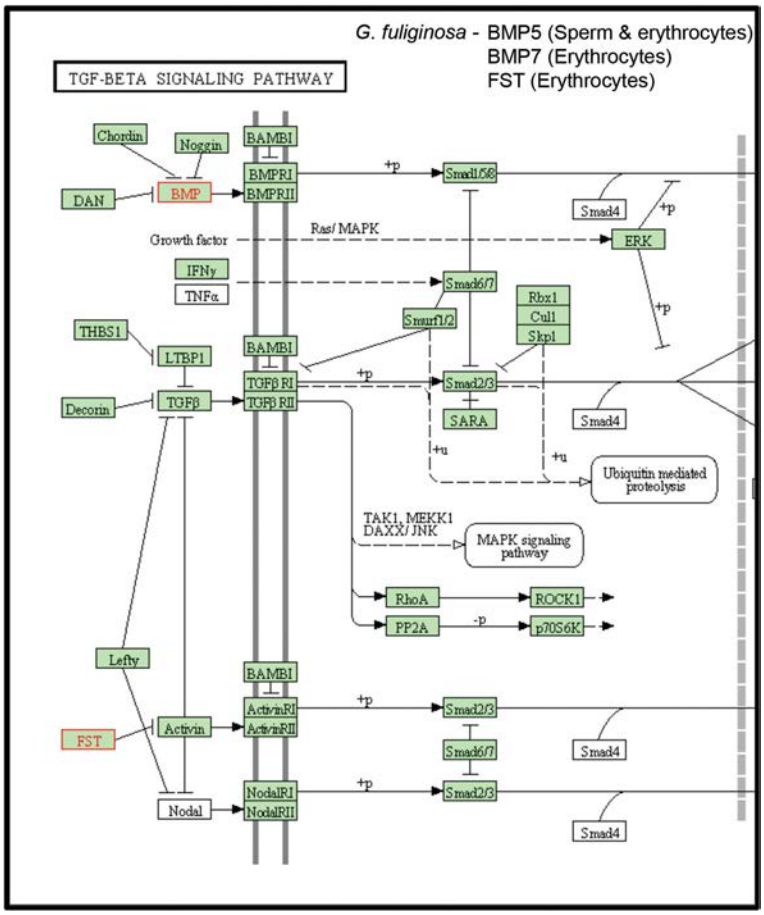


Fig. 7 TGFB/BMP pathway. Genes associated with DMR are listed and outlined in *red* in the pathway

results show that negligible large size CNV changes exist between the rural and urban populations of *G. fortis* or *G. fuliginosa*.

In contrast to our genetic results, we found a large number of epigenetic differences between urban and rural populations in both species of finches and both cell types (Fig. 2). Although DMRs were found in both species, few of the same genomic regions were differentially methylated in *G. fortis* and *G. fuliginosa*. These data suggest that methylation patterns are species-specific, even when comparing closely related species. This may mean that *G. fortis* and *G. fuliginosa* are responding to environmental changes at the urban site in different ways. The lack of overlap in DMRs between the two species may reflect differences in their diets [27]. As discussed above, dietary differences may also have contributed to the morphological differences between urban and rural populations of *G. fortis*.

Although DMRs were also found in both cell types, few of the same genomic regions were differentially methylated in sperm and erythrocytes. Because methylation is involved with cell differentiation [6, 69], some lack of similarity in erythrocyte and sperm DMR is expected. The differences between the genomic regions that were differentially methylated in sperm and erythrocytes may provide clues as to the functional significance of the epimutations. DMRs in somatic cells, such as erythrocytes, potentially reflect effects of the environment on physiology of the birds. DMRs in germ cells, such as sperm, are more likely to be transgenerationally inherited and contribute to evolution. Recent studies show that heritability of methylation variants can be high, but that this varies among loci [12]. However, without following multiple generations of individuals with known ancestry, we cannot determine which of the DMRs in our study are heritable. It is possible that many of the DMRs we detected were plastic responses to the environment. Analysis of Darwin's finches with known pedigrees - from long-term studies of banded populations - may be a way in which to distinguish heritable from non-heritable epimutations in the future.

While locations of DMRs varied between species and cell types (Fig. 4), they had genomic features in common. DMRs were typically 500–1000 bp in length (Fig. 3) and

many were clustered in 2–7 Mb regions. Most DMRs were in areas of low CpG density known as "CpG deserts" (Fig. 5). Many studies of DNA methylation have focused on the gene-silencing effects of methylation in high-density "CpG islands" near transcriptional start sites [6]. However, DMRs in other genomic regions, such as CpG deserts, can have other important effects on gene regulation and expression [6, 70]. Methylation of cytosines increases the rate of cytosine to thymine transitions [71]. Thus, over time, methylation can cause CpG-poor regions in the genome to accumulate. The persistence of conserved clusters of methylated CpG sites within CpG deserts suggests that these regions are likely conserved and under purifying selection [70]. Thus, these types of DMRs may have a functional role in regulating gene expression and could be subject to selection.

Many of the DMRs we detected were associated with metabolic and signaling genes (Fig. 6). Previous work has suggested that novel food sources at the urban site are changing the diet of finches [27]. While we did not quantify phenotypic traits related to metabolism, it is possible that DMRs associated with metabolic genes are associated with other physiological differences between the urban and rural populations.

We also found DMRs associated with genes in the bone morphogenic protein (BMP) / transferring growth factor beta (TGFß) pathway (Fig. 7). Expression of Bmp4 is related to beak shape in *Geospiza* finches [72]; however, it is unknown what factors regulate gene expression at this locus. We previously found that this pathway was differentially methylated among species of Darwin's finches [18]. These data suggest that DNA methylation may play a role in regulating expression of genes in this pathway and therefore may influence finch morphology.

Our study compared just two populations - one rural and one urban – and therefore we cannot be certain that urbanization is the key environmental change influencing finch morphology and/or epigenetics in our study. Moreover, it is possible that differences between the two populations are the result of epigenetic drift, rather than differential selection. Some dispersal of *G. fortis* between the urban and rural populations has been documented through mark-recapture studies; but it is not very common (J. Raeymaekers pers. comm.). Low levels of gene flow between populations would preclude divergence of the rural and urban populations due to drift. However, much more work is needed to understand the basis of epigenetic variation and its relationship to phenotypic variation in populations of Darwin's finches.

Conclusions

We found epigenetic differences between adjacent populations of each of two species of Darwin's finches. We do not know which of the DMRs are responses to environmental differences between the urban and rural sites, versus the result of random epigenetic drift. However, as the environmental differences between our sites are recent (<60 years) any methylation changes associated with urbanization have spread quickly. As in other recent studies [19, 20, 22], the functional relationship between environmental and epigenetic variation is not well understood. Nevertheless, these results are consistent with a potential role of epigenetic variation in rapid adaptation to changing environments. Future studies are needed to determine what effects DMR have on phenotypes, and to what extent these methylation patterns may play a role in evolution.

Additional files

Additional file 1: Figure S1. Comparison of vegetative cover at the rural site (El Garrapatero) versus urban site (Puerto Ayora, Academy Bay) over the course of the study. Cover was dervied from Normalized Difference Vegetative Index (NDVI) values generated from satellite imagery (ORNL DAAC. 2008. MODIS Collection 5 Land Products Global Subsetting and Visualization Tool. ORNL DAAC, Oak Ridge, Tennessee, USA. Accessed May 08, 2017 http://dx.doi.org/10.3334/ORNLDAAC/1241). Values range from 0-1 with 1 reprensenting the highest vegetation cover. (PDF 850 kb)

Additional file 2: Figure S2. Copy number variation (CNV) between the rural and urban populations. (A) CNV analysis summary for the *G. fortis* erythrocytes showing read depth and alignment, and CNV numbers per pool with chromosomes containing CNV indicated, and no overlap between rural and urban pools indicated. (B) CNV analysis summary for the *G. fuliginosa* erythrocytes with Read Mapping Summary, overall CNV per pool and chromosome, and no overlapping CNV identified. (PDF 20 kb)

Additional file 3: Table S1. The number of DMR detected at single window and multiple window scales at increasing levels of significance. (PDF 61 kb)

Additional file 4: Table S2. Description of multiple-window DMR detected in *G. fortis* sperm **(A)** and erythrocytes **(B)**. Description includes DMR name, chromosome number, DMR start site, length in base pair (bp), number of multiple sites, minimum p-value, CpG number per sequence length, CpG density (CpG number / 100 bp) and DMR gene association. "NA" indicates DMR associated with a gene that did not align to the zebra finch reference genome. (PDF 126 kb)

Additional file 5: Table S3. Description of multiple-window DMR detected in *G. fuliginosa* sperm **(A)** and erythrocytes **(B)**. Description includes DMR name, chromosome number, DMR start site, length in base pair (bp), number of multiple sites, minimum p-value, CpG number per sequence length, CpG density (CpG number / 100 bp) and DMR gene association. "NA" indicates DMR associated with a gene that did not align to the zebra finch reference genome. (PDF 154 kb)

Additional file 6: Table S4. Description of DMR clusters detected in *G. fortis* sperm **(A)** and erythrocytes **(B)** and *G. fuliginosa* sperm **(C)** and erythrocytes **(D)**. Description includes DMR in cluster, chromosome number, cluster start site, cluster stop site, length in bp, and minimum *p*-value. (PDF 103 kb)

Additional file 7: Table S5. Gene associations with DMR detected in *G. fortis* sperm **(A)** and erythrocytes **(B)** and *G. fuliginosa* sperm **(C)** and erythrocytes **(D)**. Description includes DMR name, gene symbol, entrez gene identification, chromosome number, start position site, ensemble gene identification number, gene description and gene classification category. (PDF 225 kb)

Additional file 8: Figure S3. MAPK signaling pathway. Genes associated with DMR are listed and outlined in red in the pathway. (PDF 109 kb)

Additional file 9: Figure S4. Glycolysis metabolism pathway. Genes associated with DMR are listed and outlined in red in the pathway. (PDF 66 kb)

Additional file 10: Figure S5. DMRs identified in pairwise comparison of pools within populations: **(A)** *G. Fuliginosa* RBC urban analysis, **(B)** *G. fuliginosa*-RBC rural analysis, **(C)** *G. fortis* RBC urban analysis, and **(D)** *G. fortis* rural analysis. Numbers indicate DMRs between urban (U) or rural (R) individual pools (1-3). "Full analysis" are DMRs identified between urban and rural pools. DMRs identified in the full analysis were found independently of within-site variation. (PDF 98 kb)

Abbreviations

BMP: bone morphogenic protein; CNV: copy number variation; DDT: dithiothreitol; DMR: differentially DNA methylated region; IP: immunoprecipitation; LMM: linear mixed effects models; NGS: next generation sequencing; PBS: Phosphate Buffered Saline; TGFß: transferring growth factor beta

Acknowledgments

We acknowledge the advice and critical review of Dr. Eric Nilsson (WSU). The following people contributed to sample collection: Céline Le Bohec, Sarah Bush, Roger Clayton, Miriam Clayton, Oliver Tiselma, Elena Arriero, Andrew Bartlow, Daniela Vargas, Emily DiBlasi, Jordan Herman, Pricilia Espina, Kiyoko Gotanda, Sofia Carvajal, Joost Raeymakers and Janaí Yepez. We thank Ms. Jayleana Barton for molecular technical assistance and Ms. Heather Johnson for assistance in preparation of the manuscript. The research was supported by a Templeton grant to MKS, National Science Foundation grants DEB-0816877 and DEB-1342600 to DHC, and an NSF Graduate Research Fellowship to SMM. Dr. Sarah A. Knutie's present address: Department of Ecology and Evolutionary Biology, Storrs, CT 06269-3043, USA. Dr. Jennifer A. H. Koop's present address: Biology Department, University of Massachusetts-Dartmouth, Dartmouth, MA 02747-2300, USA.

Funding

The research was supported by a Templeton grant to MKS, a National Science Foundation grant to DHC, and an NSF Graduate Research Fellowship to SMM.

Availability of data and materials

All the DMR sequence and genomic data obtained in the current study have been deposited in the NCBI public GEO database (GEO # GSE87825).

Author contributions

DHC and MKS designed the study; SMM, SAK and JAHK collected the samples, DB and ISR analyzed the genomic data, SMM, DHC and MKS analyzed the data and wrote the manuscript with help from the other authors. All authors read and approved the final manuscript.

Ethics approval and consent to participate

All field procedures were approved by the University of Utah Institutional Animal Care and Use Committee (protocols #07–08004, #10–07003 and #13–06010) and by the Galápagos National Park.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]Department of Biology, University of Utah, Salt Lake City, UT 84112-0840, USA. [2]Center for Reproductive Biology, School of Biological Sciences, Washington State University, Pullman, WA 99164-4236, USA.

Received: 26 January 2017 Accepted: 26 July 2017
Published online: 24 August 2017

References

1. Bossdorf O, Richards CL, Pigliucci M. Epigenetics for ecologists. Ecol Lett. 2008;11(2):106–15.
2. Day T, Bonduriansky R. A unified approach to the evolutionary consequences of genetic and nongenetic inheritance. Am Nat. 2011;178(2):E18–36.
3. Robertson M, Richards C. Non-genetic inheritance in evolutionary theory - the importance of plant studies. Non-Genetic Inherit. 2015;2:3–11.
4. Angers B, Castonguay E, Massicotte R. Environmentally induced phenotypes and DNA methylation: how to deal with unpredictable conditions until the next generation and after. Mol Ecol. 2010;19(7):1283–95.
5. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet. 2003; 33(Suppl):245–54.
6. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012;13(7):484–92.
7. Duncan EJ, Gluckman PD, Dearden PK. Epigenetics, plasticity, and evolution: how do we link epigenetic change to phenotype? J Exp Zool B Mol Dev Evol. 2014;322(4):208–20.
8. Jirtle RL, Skinner MK. Environmental epigenomics and disease susceptibility. Nat Rev Genet. 2007;8(4):253–62.
9. Crews D, Gore AC, Hsu TS, Dangleben NL, Spinetta M, Schallert T, Anway MD, Skinner MK. Transgenerational epigenetic imprints on mate preference. Proc Natl Acad Sci U S A. 2007;104(14):5942–6.
10. Richards CL, Bossdorf O, Pigliucci M. What role does heritable epigenetic variation play in phenotypic evolution? Bioscience. 2010;60:232–7.
11. Latzel V, Zhang Y, Karlsson Moritz K, Fischer M, Bossdorf O. Epigenetic variation in plant responses to defence hormones. Ann Bot. 2012;110(7):1423–8.
12. Janowitz Koch I, Clark MM, Thompson MJ, Deere-Machemer KA, Wang J, Duarte L, Gnanadesikan GE, McCoy EL, Rubbi L, Stahler DR, et al. The concerted impact of domestication and transposon insertions on methylation patterns between dogs and grey wolves. Mol Ecol. 2016;25(8):1838–55.
13. Verhoeven KJ, vonHoldt BM, Sork VL. Epigenetics in ecology and evolution: what we know and what we need to know. Mol Ecol. 2016;25(8):1631–8.
14. Skinner MK, Manikkam M, Guerrero-Bosagna C. Epigenetic transgenerational actions of environmental factors in disease etiology. Trends Endocrinol Metab. 2010;21(4):214–22.
15. Richards CL, Schrey AW, Pigliucci M. Invasion of diverse habitats by few Japanese knotweed genotypes is correlated with epigenetic differentiation. Ecol Lett. 2012;15(9):1016–25.
16. Liu QA. The impact of climate change on plant epigenomes. Trends Genet. 2013;29(9):503–5.
17. Flatscher R, Frajman B, Schonswetter P, Paun O. Environmental heterogeneity and phenotypic divergence: can heritable epigenetic variation aid speciation? Genet Res Int. 2012;2012:698421.
18. Skinner MK, Guerrero-Bosagna C, Haque MM, Nilsson EE, Koop JAH, Knutie SA, Clayton DH. Epigenetics and the evolution of Darwin's Finches. Genome Biology & Evolution. 2014;6(8):1972–89.
19. Lira-Medeiros CF, Parisod C, Fernandes RA, Mata CS, Cardoso MA, Ferreira PC. Epigenetic variation in mangrove plants occurring in contrasting natural environment. PLoS One. 2010;5(4):e10326.
20. Liu S, Sun K, Jiang T, Ho JP, Liu B, Feng J. Natural epigenetic variation in the female great roundleaf bat (Hipposideros Armiger) populations. Mol Gen Genomics. 2012;287(8):643–50.
21. Gugger PF, Fitz-Gibbon S, PellEgrini M, Sork VL. Species-wide patterns of DNA methylation variation in Quercus Lobata and their association with climate gradients. Mol Ecol. 2016;25(8):1665–80.
22. Lea AJ, Altmann J, Alberts SC, Tung J. Resource base influences genome-wide DNA methylation levels in wild baboons (Papio Cynocephalus). Mol Ecol. 2016;25(8):1681–96.
23. Zhao Y, Tang JW, Yang Z, Cao YB, Ren JL, Ben-Abu Y, Li K, Chen XQ, Du JZ, Nevo E. Adaptive methylation regulation of p53 pathway in sympatric speciation of blind mole rats, Spalax. Proc Natl Acad Sci U S A. 2016;113(8):2146–51.
24. Foust CM, Preite V, Schrey AW, Alvarez M, Robertson MH, Verhoeven KJ, Richards CL. Genetic and epigenetic differences associated with environmental gradients in replicate populations of two salt marsh perennials. Mol Ecol. 2016;25(8):1639–52.

25. Podos J. Correlated evolution of morphology and vocal signal structure in Darwin's finches. Nature. 2001;409(6817):185–8.
26. Grant PR, Grant BR. 40 years of evolution. Princeton, NJ: Princeton University Press; 2014.
27. De Leon LF, Podos J, Gardezi T, Herrel A, Hendry AP. Darwin's finches and their diet niches: the sympatric coexistence of imperfect generalists. J Evol Biol. 2014;27(6):1093–104.
28. Lamichhaney S, Berglund J, Almen MS, Maqbool K, Grabherr M, Martinez-Barrio A, Promerova M, Rubin CJ, Wang C, Zamani N, et al. Evolution of Darwin's finches and their beaks revealed by genome sequencing. Nature. 2015;518(7539):371–5.
29. Lamichhaney S, Han F, Berglund J, Wang C, Almen MS, Webster MT, Grant BR, Grant PR, Andersson L. A beak size locus in Darwin's finches facilitated character displacement during a drought. Science. 2016;352(6284):470–4.
30. Chaves JA, Cooper EA, Hendry AP, Podos J, De Leon LF, Raeymaekers JA, MacMillan WO, Uy JA. Genomic variation at the tips of the adaptive radiation of Darwin's finches. Mol Ecol. 2016;25(21):5282–95.
31. de Leon LF, Bermingham E, Podos J, Hendry AP. Divergence with gene flow as facilitated by ecological differences: within-island variation in Darwin's finches. Philos Trans R Soc Lond Ser B Biol Sci. 2010;365(1543):1041–52.
32. Sato A, O'HUigin C, Figueroa F, Grant PR, Grant BR, Tichy H, Klein J. Phylogeny of Darwin's finches as revealed by mtDNA sequences. Proc Natl Acad Sci U S A. 1999;96(9):5101–6.
33. Petren K, Grant B, Grant P. A phylogeny of Darwin's finches based on microsatellite DNA length variation. Proc R Soc Lond B. 1999;266:321–9.
34. Farrington HL, Lawson LP, Clark CM, Petren K. The evolutionary history of Darwin's finches: speciation, gene flow, and introgression in a fragmented landscape. Evolution. 2014;68(10):2932–44.
35. Hendry AP, Grant PR, Rosemary Grant B, Ford HA, Brewer MJ, Podos J. Possible human impacts on adaptive radiation: beak size bimodality in Darwin's finches. Proc Biol Sciences / R Soc. 2006;273(1596):1887–94.
36. De Leon LF, Raeymaekers JA, Bermingham E, Podos J, Herrel A, Hendry AP. Exploring possible human influences on the evolution of Darwin's finches. Evolution. 2011;65(8):2258–72.
37. Grant P. Ecology and evolution of Darwin's finches. Princeton, NJ: Princeton University Press; 1986.
38. Grant PR, Grant BR. Evolution of character displacement in Darwin's finches. Science. 2006;313(5784):224–6.
39. Kuznetsova A, Brockhoff P, Christensen R: lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0–33. 2016.
40. Bates D, Maechler M, Bolker B, Walker S. lme4: linear mixed-effects models using Eigen and S4. J Stat Softw. 2015;67:1–48.
41. Lemon J. Plotrix: a package in the red light district of R. R-News. 2006;6:8–12.
42. Champely S: pwr: Basic functions for power analysis. . R package version 12–1 2017.
43. Freeman JL, Perry GH, Feuk L, Redon R, McCarroll SA, Altshuler DM, Aburatani H, Jones KW, Tyler-Smith C, Hurles ME, et al. Copy number variation: new insights in genome diversity. Genome Res. 2006;16(8):949–61.
44. Schlotterer C, Tobler R, Kofler R, Nolte V. Sequencing pools of individuals - mining genome-wide polymorphism data without big funding. Nat Rev Genet. 2014;15(11):749–63.
45. Zhang W, Carriquiry A, Nettleton D, Dekkers JC. Pooling mRNA in microarray experiments and its effect on power. Bioinformatics. 2007;23(10):1217–24.
46. Taiwo O, Wilson GA, Morris T, Seisenberger S, Reik W, Pearce D, Beck S, Butcher LM. Methylome analysis using MeDIP-seq with low DNA concentrations. Nat Protoc. 2012;7(4):617–36.
47. Harris RA, Wang T, Coarfa C, Nagarajan RP, Hong C, Downey SL, Johnson BE, Fouse SD, Delaney A, Zhao Y, et al. Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. Nat Biotechnol. 2010;28(10):1097–105.
48. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.
49. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. Nat Methods. 2012;9(4):357–9.
50. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. Genome project data processing S: the sequence alignment/map format and SAMtools. Bioinformatics. 2009;25(16):2078–9.
51. Klambauer G, Schwarzbauer K, Mayr A, Clevert DA, Mitterecker A, Bodenhofer U, Hochreiter S. cn.MOPS: mixture of Poissons for discovering copy number variations in next-generation sequencing data with a low false discovery rate. Nucleic Acids Res. 2012;40(9):e69.
52. Lienhard M, Grimm C, Morkel M, Herwig R, Chavez L. MEDIPS: genome-wide differential coverage analysis of sequencing data derived from DNA enrichment experiments. Bioinformatics. 2014;30(2):284–6.
53. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–40.
54. Li B, Li H, Parker P, Wang J. The genome of Darwin's finch (Geospiza Fortis). GigaDB. 2012;
55. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/bioconductor package biomaRt. Nat Protoc. 2009;4(8):1184–91.
56. Cunningham F, Amode MR, Barrell D, Beal K, Billis K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fitzgerald S, et al. Ensembl 2015. Nucleic Acids Res. 2015;43(Database issue):D662–9.
57. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27–30.
58. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. 2014;42(Database issue):D199–205.
59. Skinner MK, Manikkam M, Haque MM, Zhang B, Savenkova M. Epigenetic Transgenerational inheritance of somatic Transcriptomes and epigenetic control regions. Genome Biol. 2012;13(10):R91.
60. Shnorhavorian M, Schwartz SM, Stansfeld B, Sadler-Riggleman I, Beck D, Skinner MK. Differential DNA Methylation Regions in Adult Human Sperm Following Adolescent Chemotherapy: Potential for Epigenetic Inheritance. PloS One. 2017;12(2):e0170085.
61. Grant P, Grant R. How and why species multiply: the radiation of Darwin's finches Princeton. NJ: Princeton University Press; 2008.
62. Huber SK, De Leon LF, Hendry AP, Bermingham E, Podos J. Reproductive isolation of sympatric morphs in a population of Darwin's finches. Proceedings Biological sciences / The Royal Society. 2007;274(1619):1709–14.
63. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, et al. Global variation in copy number in the human genome. Nature. 2006;444(7118):444–54.
64. Beckmann JS, Estivill X, Antonarakis SE. Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. Nat Rev Genet. 2007;8(8):639–46.
65. McCarroll SA, Kuruvilla FG, Korn JM, Cawley S, Nemesh J, Wysoker A, Shapero MH, de Bakker PI, Maller JB, Kirby A, et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. Nat Genet. 2008;40(10):1166–74.
66. Van't Hof AE, Campagne P, Rigden DJ, Yung CJ, Lingley J, Quail MA, Hall N, Darby AC, Saccheri IJ. The industrial melanism mutation in British peppered moths is a transposable element. Nature. 2016;534(7605):102–5.
67. Niu AL, Wang YQ, Zhang H, Liao CH, Wang JK, Zhang R, Che J, Su B. Rapid evolution and copy number variation of primate RHOXF2, an X-linked homeobox gene involved in male reproduction and possibly brain function. BMC Evol Biol. 2011;11:298.
68. Petren K, Grant PR, Grant BR, Keller LF. Comparative landscape genetics and the adaptive radiation of Darwin's finches: the role of peripheral isolation. Mol Ecol. 2005;14(10):2943–57.
69. Bestor TH. The DNA methyltransferases of mammals. Hum Mol Genet. 2000;9(16):2395–402.
70. Skinner MK, Guerrero-Bosagna C. Role of CpG deserts in the epigenetic Transgenerational inheritance of differential DNA methylation regions. BMC Genomics. 2014;15(1):692.
71. Cooper DN, Youssoufian H. The CpG dinucleotide and human genetic disease. Hum Genet. 1988;78(2):151–5.
72. Abzhanov A, Protas M, Grant BR, Grant PR, Tabin CJ. Bmp4 and morphological variation of beaks in Darwin's finches. Science. 2004;305(5689):1462–5.