

ARTICLE

Received 15 Jul 2014 | Accepted 5 Feb 2015 | Published 11 Mar 2015

DOI: 10.1038/ncomms7513

Epigenetic variation in the *Egfr* gene generates quantitative variation in a complex trait in ants

Sebastian Alvarado^{1,*,\dagger}, Rajendhran Rajakumar^{2,*,\dagger}, Ehab Abouheif² & Moshe Szyf¹

Complex quantitative traits, like size and behaviour, are a pervasive feature of natural populations. Quantitative trait variation is the product of both genetic and environmental factors, yet little is known about the mechanisms through which their interaction generates this variation. Epigenetic processes, such as DNA methylation, can mediate gene-by-environment interactions during development to generate discrete phenotypic variation. We therefore investigated the developmental role of DNA methylation in generating continuous size variation of workers in an ant colony, a key trait associated with division of labour. Here we show that, in the carpenter ant *Camponotus floridanus*, global (genome-wide) DNA methylation indirectly regulates quantitative methylation of the conserved cell-signalling gene *Epidermal growth factor receptor* to generate continuous size variation of workers. DNA methylation can therefore generate quantitative variation in a complex trait by quantitatively regulating the transcription of a gene. This mechanism, alongside genetic variation, may determine the phenotypic possibilities of loci for generating quantitative trait variation in natural populations.

¹Department of Pharmacology and Therapeutics, McGill University, 3655 Sir William Osler, Suite 1309, Montreal, Quebec, Canada H3G 1Y6. ²Department of Biology, McGill University, 1205 Dr Penfield, N3/6, Montreal, Quebec, Canada H3A 1B1. * These authors contributed equally to this work. † Present addresses: Department of Biology, Stanford University, 371 Serra Mall, Gilbert Biology 314, Palo Alto, California 94305-5020, USA (S.A.); Department of Molecular Genetics & Microbiology and UF Genetics Institute, University of Florida, 2033 Mowry Road, Gainesville, Florida 32610-3610, USA (R.R.). Correspondence and requests for materials should be addressed to M.S. (email: moshe.szyf@mcgill.ca) or to E.A. (email: ehab.abouheif@mcgill.ca).

Understanding how variation in quantitative traits is generated is important for mapping the genetic basis of disease, improving plant and animal breeding, and predicting evolutionary changes^{1,2}. The field of quantitative genetics has produced a wealth of knowledge regarding the genetic basis of quantitative trait variation, including the number, distribution and identification of large and small effect loci underlying quantitative traits^{1–3}. Little is known, however, about how the environment interacts with such loci and whether this interaction can generate quantitative trait variation^{4,5}. Several studies have shown that epigenetic mechanisms, such as DNA methylation and histone modifications, can mediate gene-by-environment interactions during development to generate discrete phenotypic variation^{6–9}. For example, in rats, DNA methylation mediates the onset of stress tolerance in response to the presence/absence of postnatal maternal care⁸, and in social insects, DNA methylation plays a role in regulating discrete morphological⁷ and behavioural¹⁰ differences between queen and worker castes. It remains unclear, however, what role DNA methylation plays in generating quantitative trait variation.

Here, we use ant societies to address this question because continuous variation in the size of individuals in the worker caste is a key trait associated with the division of labour, where differently sized workers can specialize in tasks such as excavation, brood transport and foraging^{11–13}. All holometabolous insects (ants, bees, wasps, beetles, butterflies, moths and flies) do not grow as adults^{14,15}. All growth occurs during the larval phase when larvae molt from one instar to the next and during each instar^{14,15}. Adults, due to their tough inflexible exoskeleton, do not molt and therefore their final body size is determined entirely by the end of larval development just before metamorphosis^{14,15}. The final size of a worker in an ant colony is therefore established during development¹⁶. Determination of final size of a worker is polyphenic, which means that the same genome can produce a spectrum of final adult sizes in response to environmental factors^{11,16,17}. Studies in the wild and in the lab spanning more than 100 years have identified nutrition levels and social interactions as critical environmental factors that influence the final size of workers during larval development^{16–19}. In ant species where genetic variation in the worker caste has been shown to be high, worker larvae in the colony remain polyphenic, but may vary quantitatively in their response to these environmental factors^{20–22}.

In the ant genus *Camponotus*, variation in worker size has been shown to be influenced by nutrition in the form of protein, vitamins and minerals^{23,24}. We chose the ant species *C. floridanus* because worker size variation is continuous, genetic variation between workers in a *C. floridanus* colony is low (workers are on average 75% related)²⁵ and extensive sequencing efforts indicate that within a colony there is no allelic bias between differently sized workers²⁶. Therefore, genetic variation alone cannot fully explain the diversity of worker size found in *C. floridanus*. Finally, the genome of *C. floridanus* is sequenced and has a DNA methylation system^{26,27}, thereby making this species an excellent model for understanding how gene-by-environment interactions generate quantitative trait variation in natural populations.

DNA methylation in vertebrates occurs in regulatory elements that reside in promoter regions as well as inside and outside of gene bodies throughout the genome^{28,29}, and is known to repress gene function through the covalent modification of cytosine residues within these elements³⁰. In social insects, however, DNA methylation occurs primarily in gene bodies^{28,29,31} and is generally associated with caste-specific alternative splicing^{26,32,33} and gene expression^{7,32}. There is functional evidence in honeybees that gene body methylation regulates alternative

splicing³⁴ and can repress gene expression and phenotype⁷ in a caste-specific manner. Knockdown of a DNA methylating enzyme (*Dnmt3*) in adult honeybee fat tissue affects alternative splicing³⁴, whereas knockdown of this enzyme in developing honeybee larvae results in the appearance of queen-like morphologies and a change in transcription levels of several types of genes, including those involved in growth and metabolism⁷. One of these growth-regulating genes *Target of rapamycin* (*Tor*) is more methylated³³ and is expressed at lower levels³⁵ in developing workers than in queens. Knockdown of *tor* in queen-destined larvae results in adults with worker characteristics³⁵, indicating that gene body methylation is involved in repressing the expression of *Tor* in workers.

In this study, we examine the hypothesis that naturally occurring inter-individual differences in a quantitative trait can also be generated by environmental variation through quantitative inter-individual differences in DNA methylation. We show that in colonies of the ant *C. floridanus*, natural inter-individual variation in the size of workers correlates with the natural inter-individual variation in the state of methylation of the highly conserved gene *Egfr*. We performed functional experiments to demonstrate that genome-wide DNA methylation indirectly regulates the quantitative methylation of *Egfr* to generate quantitative variation in the size of workers. By linking a continuous distribution in a trait in a natural population with a continuous distribution of DNA methylation states in a single gene, our study provides an epigenetic mechanism for generating quantitative variation in organismal phenotypes.

Results

Worker size variation and development in natural populations.

To address the role that DNA methylation plays in the regulation of worker size variation in *C. floridanus*, we first determined the size distribution of adult workers. We found that it is continuous with two peaks in frequency that represent the two worker subcastes—minor ‘m’ and major ‘M’ workers¹¹ in Fig. 1a,b. Although all workers form a continuous size distribution, minors and majors are distinguished by head allometry (the size of their heads relative to their bodies; Fig. 1c). To identify the developmental stage where this dramatic variation in worker size is established, we determined the number of larval instars. We found that there are a total of four larval instars and that most of the growth differences in size between workers are established during the 4th (final) instar (Fig. 1d). The early 4th instar larvae have not yet experienced this rapid growth (‘early’ in Fig. 1d), whereas late 4th instar larvae undergo a burst of growth producing a continuous range of final larval sizes (Fig. 1d). Because final adult size is determined at the end of larval development in insects^{14,15}, including ants¹⁶, we can therefore infer that the smaller larvae (‘m’ in Fig. 1d) will develop into adult minor workers, whereas the larger larvae (‘M’ in Fig. 1d) will develop into major workers.

Worker size variation is linked to DNA methylation. We then tested if differential methylation at the genomic and gene levels during development may be involved in generating the continuous distribution of worker size in *C. floridanus*. Before examining the entire size continuum, however, we first used the extreme ends of the distribution to screen for dynamic changes in global DNA methylation that are associated with the early and late phases of larval growth during the 4th instar. If DNA methylation plays a role in regulating worker size variation, then we expected to find higher levels of methylation in the smallest worker larvae and lower levels of methylation in the largest worker larvae. Indeed, we found that early 4th instar larvae have

significantly lower levels of global methylation relative to late 4th instar larvae (Fig. 2a) and 4th instar minor worker larvae were significantly hypermethylated compared with 4th instar major worker larvae (Fig. 2a). We next compared expression

levels of key enzymatic regulators of DNA methylation⁹ (DNA methyltransferase 1 (*Dnmt1*), DNA methyltransferase 3 (*Dnmt3*), Ten-eleven-translocation 2 (*Tet2*), Methyl-binding domain (*Mbd*) and Methyl CpG-binding protein 2 (*Mecp2*) and histone modification⁹ (Histone deacetylase 1 (*Hdac1*), Histone deacetylase 3 (*Hdac3*), Histone acetyltransferase (*Hat*) and Lysine-specific demethylase 1 (*Lsd1*)) and asked whether alterations in the level of expression of these regulators correspond to differences in global DNA methylation. We found that differences in levels of gene expression of *Dnmt1*, *Dnmt3*, *Tet2*, *Mbd* and *Mecp2* (Fig. 2b,c and Supplementary Fig. 1a–d) as well as *Hdac1* and *Hat* (Supplementary Fig. 1d,f) correspond to differences in global methylation levels during the 4th instar. Considering that key enzymatic regulators of DNA methylation, *Mbd* and *Mecp2*, are known to interact with HDACs³⁶, both regulators of DNA methylation and histone modification may contribute to the pattern of global DNA methylation. Our results therefore suggest a link between global DNA methylation and the regulation of the continuous worker size variation during the fourth instar in *C. floridanus*.

DNA methylation regulates worker size variation. To determine whether global DNA methylation plays a functional role during development to regulate the continuous size variation of workers, we manipulated levels of global methylation during the early 4th instar. We used a hypomethylating agent, the inhibitor of DNA methyltransferases 5-AZA-dCytidine (5-AZA-dC)³⁷, and a hypermethylating agent, the methyl donor S-adenosyl methionine³⁸ (SAM). Relative to controls, we found that 5-AZA-dC significantly reduced genomic methylation (Supplementary Fig. 3a) and shifted the continuous size distribution by significantly increasing the mean size of adult workers (Fig. 3b,e,f, Supplementary Fig. 2a–c). In contrast, we found that SAM increased genomic methylation as expected (although not statistically significant; Supplementary Fig. 3b) and shifts the continuous size distribution by significantly decreasing the mean size of adult workers (Fig. 3a,c,d and Supplementary Fig. 2d–f). Although SAM has many biological functions, its role in mediating global DNA methylation and sizing is reinforced by the contrasting effect on genomic methylation levels and phenotype caused by 5-AZA-dC, a well-known hypomethylating agent. Finally, because worker size is sensitive to environmental conditions (nutrition and social interactions), the size range of workers emerging from our control replicates is always smaller than in our lab-reared wild-type colonies (see insets Fig. 3a,b). This means that we are generally underestimating the magnitude of the effect of these drug treatments on shifting the continuous size distribution of workers (Figs 1a and 3a,b). These results demonstrate that, during larval development, global DNA

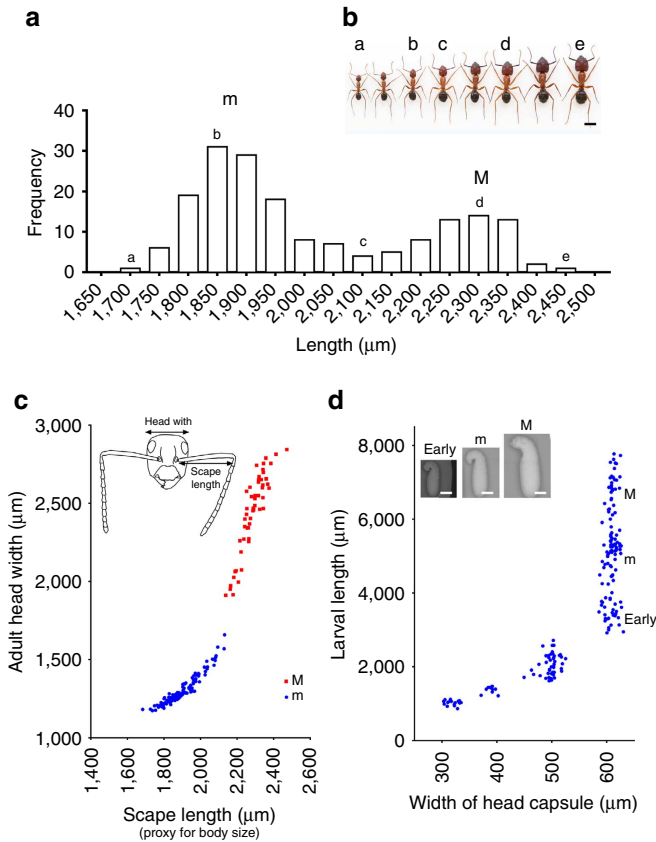


Figure 1 | Continuous size distribution of workers in *C. floridanus*. (a) Continuous distribution of adult size of *C. floridanus* workers ($n = 179$). This distribution is based on scape length (b), a common proxy for body size. (b) Representative categories of sizing continuum in natural *C. floridanus* populations that exists from smallest a to largest e. Magnification is $\times 8$. (c) Allometric classification of adult *C. floridanus* workers into major and minor workers. Adult head width was plotted against adult scape length. Minor and major adult workers are distinguished by a break in the allometry curve. Blue circles indicate minor workers, red squares indicate major workers. (d) Developmental stages and establishment of final worker sizes ($n = 197$); minor worker-destined larvae are indicated by m and major worker-destined larvae are indicated by M. Magnification of inset is $\times 30$.

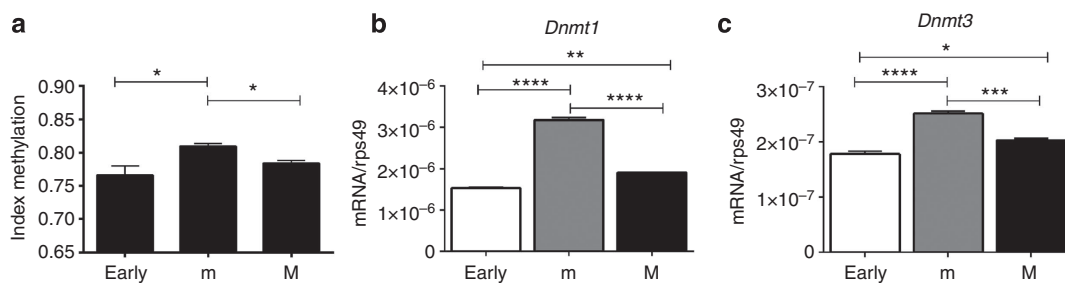


Figure 2 | DNA methylation profiles during development. (a) Genomic methylation in early 4th instar, as well as minor and major larvae in late 4th instar. Genomic methylation is shown as a ratio of [MspI/HpaII] cuts to provide an index of overall DNA methylation. Transcription level of (b) *Dnmt1* and (c) *Dnmt3* in early 4th instar larvae (white bar), late minor worker larvae (grey bar) and late major worker larvae (black bar). Bars indicate mean, whereas error bars indicate \pm s.e.m. Statistical significance values for Student's *t*-test are as follows: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$. Sample sizes for each larval category is $n \geq 30$ individual larvae.

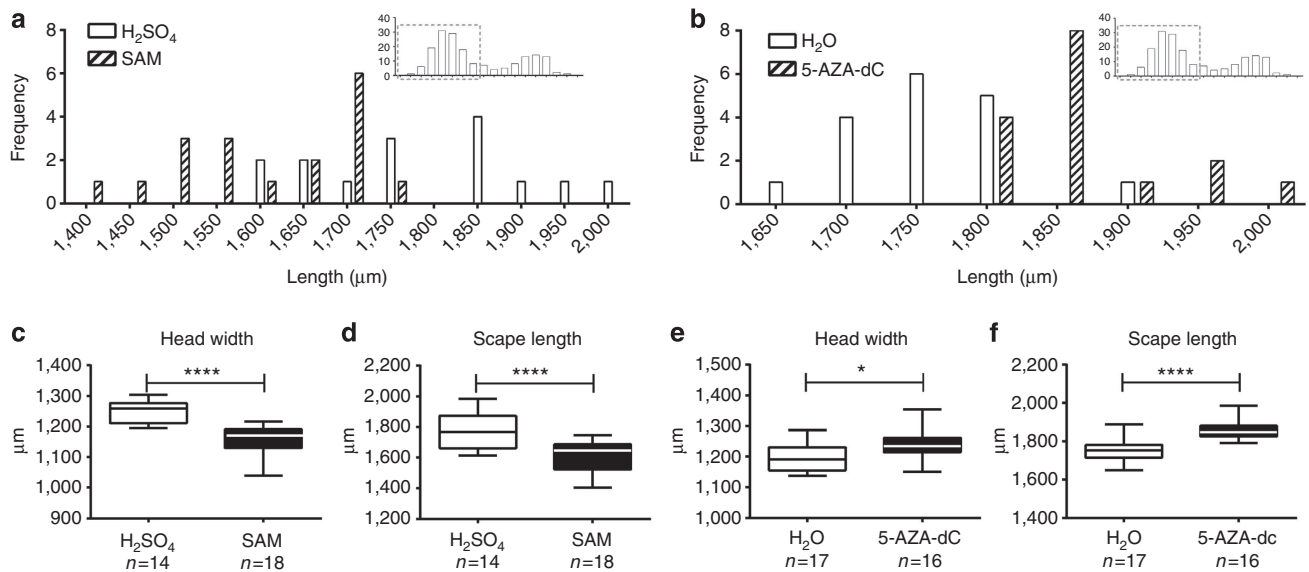


Figure 3 | Pharmacological manipulation of global (genome-wide) DNA methylation shifts mean of continuous worker size distribution. Hatched bars indicate effect of (a) SAM ($n=18$) and (b) 5-AZA-dC ($n=16$) administration at the early 4th instar, whereas white bars indicate controls (H_2SO_4 for SAM and H_2O for 5-AZA-dC) relative to WT distribution of sizes seen in inset. SAM-treated animals (black box) compared with control (white box; $n=14$) for (c) head width and (d) scape length. 5-AZA-dC-treated (black box) animals compared with control (white box; $n=17$) for (e) head width and (f) scape length. Boxplot whiskers indicate min and max. Box defined by 25th percentile, mode and 75th percentile. Bars indicate mean and error bars indicate \pm s.e.m. Statistical significance values for Student's t -test are as follows: * $P<0.05$, **** $P<0.0001$. Note that all scape length and head width measurement data are normally distributed (tested with Shapiro-Wilk's test).

methylation is involved in regulating continuous worker size distribution in *C. floridanus*.

EGFR regulates worker size variation. To identify potential targets of DNA methylation, we assayed the expression of a panel of genes (Juvenile hormone acid methyltransferase (*JHMT2*), Juvenile hormone esterase (*JHE2*), Target of rapamycin (*Tor*), Phosphatase and tensin homologue (*Pten*), *chico*, Epidermal growth factor receptor (*Egfr*)) from pathways known to play critical roles in regulating discrete size differences in social insects (Supplementary Fig. 4)^{35,39–41}. We found that *Egfr* shows the most dramatic increase in expression within 4th instar minor worker larvae (Fig. 4a and Supplementary Fig. 4f). EGFR acts as a cell surface receptor that binds specific extracellular protein ligands, including growth factors from the epidermal growth factor and transforming growth factor (alpha) families. Once bound, EGFR stimulates phosphorylation and activates downstream pathways, including the Mitogen-activated protein kinase, Serine-threonine protein kinase and c-Jun N-terminal kinase pathways, leading to the regulation of DNA synthesis, growth, cell proliferation, differentiation and many other vital cellular processes⁴². In honeybees, dietary cues (royal jelly) stimulate EGFR signalling, which regulates size and developmental timing, leading to the development of queens³⁹. Furthermore, in fruit flies, mapping loci that underlie quantitative traits (quantitative trait loci or QTLs) has identified *Egfr* as a genetic locus with a major effect on variation in quantitative sizing⁴³. To determine the developmental role of EGFR in regulating the continuous variation in worker size in *C. floridanus*, we inhibited EGFR signalling during the early 4th instar using a pharmacological inhibitor (PD 153035). This inhibitor is known to be specific to EGFR⁴⁴, and as expected, significantly inhibits EGFR phosphorylation 24h post treatment (Supplementary Fig. 3c,d). We discovered that, relative to controls, inhibition of EGFR shifts the continuous size distribution by significantly increasing the mean size of adult

workers (Fig. 4c–e and Supplementary Fig. 2g–i). This dramatic effect of EGFR inhibition indicates that EGFR is involved in regulating the continuous size distribution of workers in *C. floridanus*.

DNA methylation regulates worker size variation through EGFR.

To determine whether global DNA methylation developmentally regulates continuous size distribution in workers through EGFR, we assayed *Egfr* methylation and expression levels after we experimentally treated early 4th instar larvae with SAM and 5-AZA-dC. We found an inverse relationship between levels of global DNA methylation and *Egfr* expression. Increasing global methylation by SAM treatment during the 4th larval instar results in decreased methylation and increased expression of *Egfr* (Supplementary Fig. 5a,b), whereas decreasing global methylation by 5-AZA-dC treatment results in increased methylation and decreased expression of *Egfr* (Supplementary Fig. 5c,d). This indicates that other intermediate genes or processes downstream to global changes in DNA methylation state mediate this inverse relationship^{45,46}. Furthermore, DNA methylation and EGFR signalling may regulate worker size differences by affecting developmental timing. In ants, increasing the duration of development gives worker larvae more time to grow resulting in larger adult workers, whereas decreasing the duration of development gives larvae less time to grow resulting in smaller adult workers⁴⁷. We found that the duration of development was shortened following SAM treatment, whereas it was extended after EGFR inhibition (Supplementary Fig. 6a,c). 5-AZA-dC did not affect developmental timing possibly because higher concentration may be required to elicit such a response or because 5-AZA-dC affects on sizing may operate through a developmental timing-independent mechanism. Our results indicate that global DNA methylation indirectly regulates the expression and methylation levels of *Egfr* during development, which in turn regulates the continuous size distribution of workers possibly through changes in developmental timing.

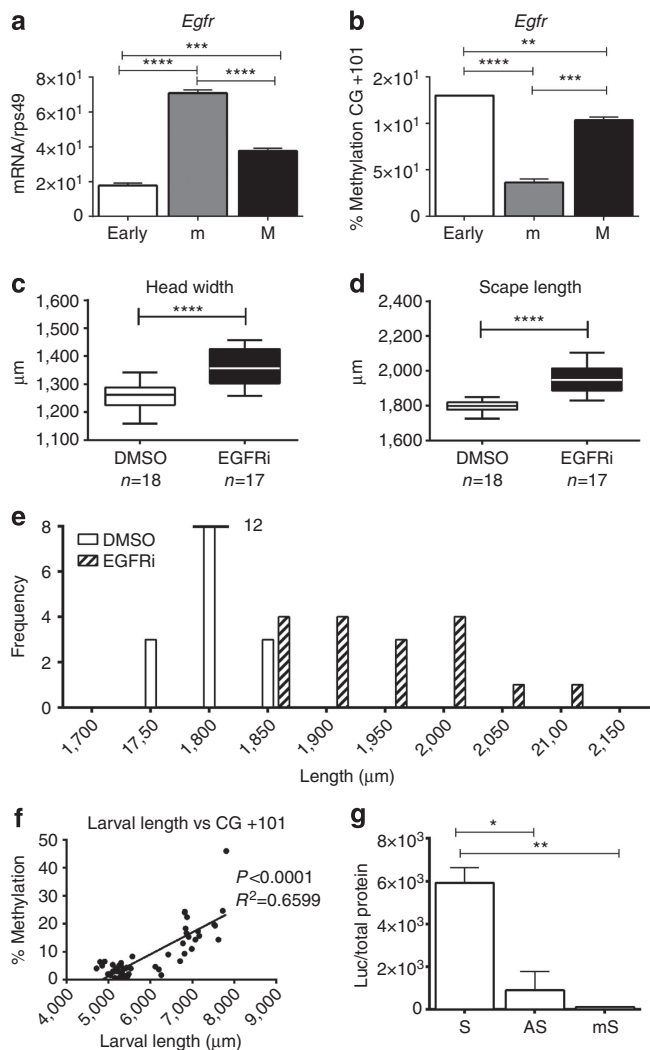


Figure 4 | The role and regulation of *Egfr* in generating quantitative variation in worker size. (a) *Egfr* expression in early 4th (white bar), late 4th instar minor (grey bar) and major larvae (black bar; $n \geq 30$). (b) Percent (%) methylation of site CpG +101 in *Egfr* for early 4th instar (white bar), late 4th instar minor (grey bar) and major larvae (black bar; $n \geq 30$). EGFR inhibition of early 4th instar worker larvae (black box) compared with control DMSO (white box) for (c) head width and (d) scape length. Scape and head width measurement data are normally distributed (tested with Shapiro-Wilk's test). Boxplot whiskers indicate min and max. Box defined by 25th percentile, mode and 75th percentile. (e) EGFR inhibition at the early 4th instar and its effect on sizing distribution (hatched bars; $n = 17$) compared with control DMSO (white bars; $n = 18$) relative to WT distribution of sizes seen in inset. (f) Linear regression between percent (%) methylation of *Egfr* site +101 (y-axis) and final larval length in μm (x-axis) across a spectrum of larval sizes ($n = 50$). (g) *Egfr* promoter cloned into luciferase reporter pCpGI in sense (S), antisense (AS) and methylated sense (mS). Bars indicate luciferase activity from the pCpGI construct relative to total protein ($n = 3$). Data shown as mean and error bars indicate \pm s.e.m. Statistical significance values for Student's *t*-test are as follows: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$.

Quantitative methylation of *Egfr* generates worker size variation. We then discovered that DNA methylation through the specific methylation of *Egfr* not only regulates, but also generates a large part of the continuous worker size distribution. We found that quantitative differences in DNA methylation of *Egfr* is highly correlated to the final size of individual worker larvae, which

spans the entire larval size range of the distribution, including minor and major worker larvae. Because final adult size is determined at the end of larval development^{14–16}, we can therefore infer that these quantitative differences in DNA methylation of *Egfr* are also correlated to final adult size. In the *C. floridanus* genome, CpG dinucleotide methylation is primarily concentrated at the beginning of the protein-coding region of all genes²⁶. We therefore screened the first 225 bp of *Egfr* and identified all CpG dinucleotides. We then determined the % methylation for each of these CpG dinucleotides in 50 individual late 4th instar larvae that represent the entire size continuum of workers. Of all the CpG dinucleotides screened, we discovered four sites for which the level of DNA methylation is significantly correlated to the final size of late 4th instar larvae (Fig. 4b,f and Supplementary Fig. 7). Two of these sites (CG +101, $R^2 = 0.66$ and CG +182, $R^2 = 0.24$) remained statistically significant after Bonferroni correction for multiple comparisons (Supplementary Fig. 7 and Supplementary Table 2). We then cloned this sequence into a CpG-free luciferase construct to determine whether these CpG dinucleotides, when methylated, affect the transcription of *Egfr*. We show that this differentially methylated region of *Egfr* could direct the transcription of this luciferase reporter in human HEK293 cells suggesting that it is a *bona fide* transcription regulatory region and that DNA methylation can indeed repress its transcriptional activity (Fig. 4g). Finally, in order to rule out sequence variation in this region of *Egfr* as a possible cause for the observed size differences of individual workers, we sequenced this region from the smallest (minor) and the largest (major) workers and found no genetic differences (Supplementary Fig. 8). Collectively, our results indicate that DNA methylation quantitatively regulates *Egfr* during larval development to generate a large part of the continuous worker size distribution.

Discussion

Generation of continuous worker size variation by the epigenetic control of *Egfr* can be hypothesized as follows (Fig. 5): variation in environmental factors leads to the hypermethylation (increase) or hypomethylation (decrease) of genomic DNA methylation levels during development of final instar larvae. Although our study did not determine the specific environmental factors that cause natural variation in DNA methylation, previous studies in ants, including other species in the genus *Camponotus*, have established both nutritional variation and social interactions as causes for variations in worker size^{16–19,23,24}. Hyper- or Hypomethylation in genomic DNA methylation levels then translates into the differential quantitative methylation of *Egfr* through intermediary genes or processes. Histone modifications are known to be key regulators in *C. floridanus*⁴⁸, making them a possible candidate for this intermediary regulation of the inverse relationship between global DNA methylation and methylation of *Egfr*. This results in quantitative inter-individual differences in the transcription of *Egfr* in the colony, which in turn generates the differential growth attained by late final instar larvae. In *C. floridanus*, we show that quantitative methylation at *Egfr* explains a large proportion of the variance (>65%) of final larval size in natural colonies. This, combined with our drug treatment results and the fact that final adult size is determined by the final size of larvae^{14,15}, suggests a causal relationship between methylation, *Egfr*, and the final size of adult workers.

In ants, *Egfr* may be positioned as a key regulator of conserved pathways; in honeybees, diet activates EGFR to regulate several important downstream pathways, including the TOR, juvenile hormone and insulin signalling pathways³⁹. In other animals, EGFR signalling is also known to play a role in regulating DNA methylation through downstream activation of DNMTs⁴⁹

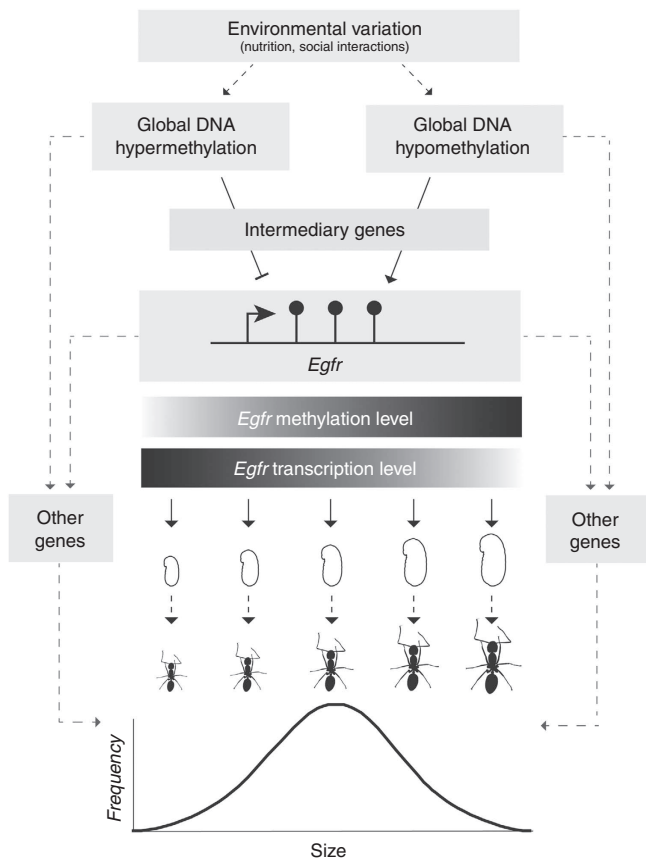


Figure 5 | Model for the generation of a continuous distribution of worker size through quantitative DNA methylation of *Egfr*. Solid black lines indicate our results, whereas dashed black lines indicate our inferences from the literature. Lines with perpendicular tips indicate repression, whereas lines with arrowheads indicate activation. Environmental variation may lead to either global DNA hyper- or hypo-methylation (dashed black lines). Intermediary genes or processes quantitatively modulate this global DNA methylation state (solid black lines) into a specific level of *Egfr* methylation (black lollipops and box with increasing gradient from left to right, where an increase in black corresponds to increasing levels of methylation), which results in a specific level of *Egfr* transcription level (box with decreasing gradient from left to right, where a decrease in black corresponds to decrease in *Egfr* transcription levels). Other genes may also be involved in generating quantitative size variation (dashed black lines). The level of *Egfr* transcription then translates to quantitative size variation of worker larvae (solid black lines). Finally, we infer that the continuous size distribution of final size worker larvae translates into the continuous size distribution of adult workers (dashed black lines).

resulting in altered cellular growth. This raises the possibility that in ants, *Egfr* may not only be a target of DNA methylation but may in turn also regulate DNA methylation. It is important to note that the main goal of our study was to understand the developmental role of DNA methylation in producing quantitative variation in worker size and not caste determination, that is, producing minor and major worker subcastes with their discrete allometric differences between head and body size (Fig. 1c). Caste determination is a complex process that may require more than just a size increase through *Egfr*. Major worker subcastes have many discrete features other than size¹¹, which are not present in minor worker subcastes, and may require juvenile hormone and other factors.

The quantitative methylation of specific loci, like *Egfr*, may represent a more general mechanism for the generation and evolution of quantitative trait variation across animals. DNA methylation, the EGFR pathway and specific pathways downstream of EGFR are highly conserved across animals, and DNA methylation is known to be transgenerationally inherited^{9,50–52}. In vertebrates, quantitative differences in methylation of an inserted retroviral element in front of the *Agouti* gene in mouse defines differences in coat colour⁵³, and the distribution of coat colours could be shifted by altering the methyl content in maternal diet⁵³. It remains to be shown, however, whether natural inter-individual epigenetic variation of this inserted retroviral element generates variation in coat colour in natural populations. This mechanism may also underlie associations found between epigenetic variation and disease susceptibility, like that found for Type II diabetes in humans⁵⁴, as well as quantitative traits, like that found for flowering time and height in plants⁵¹.

Finally, our results hold important implications for quantitative genetics. The eventual unification of Mendelian inheritance with Darwin's theory of natural selection was made possible by the infinitesimal model⁵⁵, which assumes that quantitative trait variation is generated by the action of an infinite number of loci that have small and equal effects on the phenotype^{3,56}. The evolution of quantitative traits is therefore thought to occur through random mutations across these loci^{3,56}. However, the empirical search for QTLs has revealed that trait variation often maps to specific genetic regions of small or large effect, and with specific functions^{4,57}. The apparent gap between the assumptions of the infinitesimal model and the results of QTL analyses is further exacerbated by the fact that countless studies have demonstrated that QTLs cannot in themselves explain all heritable variation underlying quantitative traits⁵, such as growth or size in humans⁵⁸, *Arabidopsis*⁵⁹ and yeast⁶⁰. These difficulties underscore the many challenges that remain in understanding the genetic basis of quantitative trait variation^{3,5}. Our findings may help resolve these outstanding challenges in the field of quantitative genetics because they show that the phenotypic possibilities of a genetic locus may be determined by both genetic variation and the influence of the environment through quantitative DNA methylation. Therefore, alongside genetic variation, the environment can generate a vast array of quantitative trait variation in natural populations.

Methods

Collection of samples. We collected mated queens from Tallahassee, Florida, USA. Mature colonies originating from single queens were maintained in plastic boxes with glass test tubes filled with water constrained by cotton wool, and were fed a combination of mealworms, crickets, fruit flies and Bhatkar-Whitcomb diet as in Rajakumar *et al.*¹⁶ All colonies were maintained at 27 °C, 70% humidity and 12 h day:night cycle. Larval and adult sampling, genotyping and pharmacological experiments were conducted on a single colony to control for colony and population variation and began after the colony had matured for approximately 4 years. In order to conduct SAM (luminometric methylation assay (LUMA)), 5-AZA-dC (LUMA) and EGFRi (western blot) validation assays, which are described in more detail below, we pooled samples to be treated from multiple colonies because the assays required a significant amount of material.

Determination of the number of instars for worker larvae. To describe the epigenetic status of specific developmental stages of larval development, the number of larval instars was determined. The colony which larvae were taken from was not in the process of producing reproductives (males or queens). This is important because larvae of *Camponotus* species that become reproductive have been suggested to have a different number of instars compared with worker larvae⁶¹. To discriminate between instars, widths of the head capsules of larvae were measured as previously described⁶². Larvae analysed that were designated as minor 'm' or major 'M' worker larvae are terminal larvae, meaning that they have reached the end of larval development, which we have determined is at the end of the fourth instar. m and M worker larvae share the same head capsule size but differ in overall body size. A larvae is classified as terminal if it meets the two following requirements: the gut of the larvae is completely black and the fat cells of

the larvae are at peak density and completely engulf the gut⁶³. When the larvae exhibit those two characters, then they are at their terminal stage. Larvae designated m or M that were used specifically for analysis (LUMA, bisulfite sequencing, quantitative-PCR) were the smallest and largest terminal larvae, respectively, in order to ensure their fate. Adults described as m or M is designated as such based on allometric differences (detailed in 'Measuring of adult workers' section).

Culturing of pharmacologically treated larvae. To determine the role of methylation during worker larval development, early 4th instar (the last instar based on Fig. 1d) larvae were selected for methylome and EGFR signalling manipulation. SAM, a methyl donor to DNA methyltransferases³⁸ (B9003S, New England Biolabs), was applied at a concentration of 32 mM (dissolved in 10% ethanol and 0.005 M sulfuric acid). 5-AZA-dC, a cytidine analogue inhibitor of DNA methylation³⁷ (A3656, Sigma-Aldrich), was applied at a concentration of 10 mM (dissolved in water). Last, the highly specific and potent quinazoline EGFR inhibitor, PD153035, which acts through the suppression of epidermal growth factor-dependent EGFR phosphorylation⁴⁴ (234491, EMD Millipore), was applied at a concentration of 10 μ M (dissolved in dimethylsulphoxide). Larvae were isolated and placed laterally on a Petri dish with the aid of a microscope. A volume of 4 μ l of solution was then applied topically allowing for absorption and larval feeding. After treatment, larvae were setup in plastic boxes and fed in the same manner as colonies. For each treatment, 40 larvae were treated and 80 adult minor workers were supplemented to the box to care for them. Half of the larvae were collected at the end of larval development for further quantitative gene expression as well as pyrosequencing, while the other half were left to develop. Timing of eclosion (the beginning of adulthood) was monitored for the remaining larvae of each treatment. Newly emerged adults were subsequently measured to detect any morphological effects of the drug treatments. To validate effects of SAM and 5-AZA-dC on global methylation levels, 50 larvae treated with SAM, 5-AZA-dC and their respective controls were collected for LUMA.

Measuring of adult workers. Workers were measured for several parameters for the purpose of identifying any shifts in size that might have been caused by the manipulation of larval methylation or EGFR signalling inhibition. In particular, we measured scape length, head width, thorax length, thorax width and mandible length as in Diniz-Filho (1994). In *Camponotus*, Diniz-Filho (1994) found that head width exhibited positive allometry (Fig. 1c), whereas scape length was isometric and suggested that for a bivariate analysis of allometry, scape length could be used as an independent variable (due to its isometry)⁶⁴. Isometric measures, like scape length, are highly correlated and proportional to body size⁶⁴ and therefore serve as independent and accurate proxies for body size to examine the size frequency of individuals of a colony. In a similar manner, head width and scape length were used in another *Camponotus* species⁶⁵ as well as other polymorphic ant species¹¹. An added advantage for measuring the length of the scape is its simplicity and therefore it reduces any technical variation in measurements between individuals. Therefore, to determine the relative distribution of sizes (Fig. 1a,b), we used scape length as a proxy for body size. By characterizing the allometry found within the continuous size distribution of workers, we could classify minor (m) and major (M) worker adults (Fig. 1a,c).

Microscopy. We used a Zeiss Discovery V12 stereomicroscope and Zeiss Axiovision software to measure the larvae (in μ m) and adults (in μ m). For larval imaging, we used an Olympus TM3000 tabletop scanning electron microscope.

DNA and RNA extraction. Larvae were collected and pooled ($n = 20$) for each instar and immediately frozen at -80°C . Since our trait of interest, size, is a variable specific to a large proportion of cells within the animal, we assumed that heterogeneous tissues across individual larvae would still provide data relevant to size as a trait. DNA and RNA were both extracted using Allprep DNA/RNA extraction kits (80204, Qiagen) as instructed for animal tissues. Homogenization of samples was achieved using RLT + provided with pestle grinding. Individual larvae selected at the fourth instar were also processed in a similar manner. A DNase on-column step was done in all samples during RNA extraction. All samples were quantified using Nanodrop 1000 (Thermo Scientific).

Luminometric methylation assay. LUMA is a high-throughput assay used to determine global (genome-wide) DNA methylation. The LUMA method used in our study is a modification described by Karimi *et al.*⁶⁶ LUMA involves the digestion of genomic DNA by a methylation-sensitive (HPAII) or -insensitive (MspI) restriction enzymes in combination with an internal control restriction enzyme (*EcoRI*) to normalize the DNA input. *EcoRI* (FD0274), *HpaII* (FD0514) and *MspI* (FD0541) were all purchased from Thermo Scientific. Both *HpaII* and *MspI* restriction enzymes recognize and cleave 5'-CCGG-3' sequences producing 5'-CG overhangs, whereas *EcoRI* recognizes and cleaves 5'-GAATTC-3' sequences and produces 5'-AATT overhangs. The extent of cleavage is determined by a bioluminescent polymerase extension assay, which measures the filling in of the

sticky ends generated by the enzymatic digestion using a four-step pyrosequencing reaction.

Samples were incubated (37°C , 4 h) and then heat inactivated (80°C , 20 min). Digested genomic DNA (15 μ l) was mixed with pyrosequencing annealing buffer (15 μ l; Qiagen). Samples were transferred to 24-well pyrosequencing plates for sequencing (PyroMark 24; Biotage). Peak heights for C and A represent the *HpaII* and *MspI* cuts (methylation) and *EcoRI* (input DNA), respectively. The formula to calculate an index of genomic methylation is: $[\text{MspI (C)/ EcoRI (A)}]/[\text{HpaII (C)/ EcoRI (A)}]$. The higher the ratio between *MspI* cuts (cleaves all CCGG sites) and *HpaII* cuts (cleaves only unmethylated CCGG), the higher the methylation level. All samples were run in triplicate. This method only measures an absolute index of methylation across the entire genome so equal bidirectional differences in DNA methylation are unlikely to register observable differences.

Bisulfite mapping, sequencing and expression analyses. DNA was treated with sodium bisulfite and primers (Supplementary Table 1) were designed for converted products of the 5' region (based on Kamakura, 2011) of *C. floridanus Egfr* (scaffold 550: bp113394-bp113619; located using GBrowse of the Hymenoptera Genome Database⁶⁷ Bankit ID KP325210). Bisulfite conversion was done with Epitect Bisulfite Conversion kit (Qiagen). Bisulfite PCRs were amplified using two rounds of PCR using outer and nested primers (see Supplementary Table 1). Cycling conditions involved an initial step of 5 min at 95°C followed by 35 cycles of 95°C for 1 min, T_m for 2.5 min, 72°C for 1 min and followed by 5 min of 72°C . PCR products were sequenced using the Biotage Pyrosequencer as previously described⁶⁸.

For all samples, 500 ng of RNA was subjected to RT-PCR according to manufacturer's protocols (Roche) and quantified using quantitative PCR on the Lightcycler 480 (Roche). Primers for all genes (Supplementary Table 1) were created across exon boundaries. As a housekeeping gene, we used *RPS49* for normalization as it was previously shown to have stable expression both across larval development and following juvenile hormone manipulation in *Apis mellifera*⁶⁹. Quantitative PCR was amplified with a pre-incubation at 95°C for 10 min followed by 45 cycles of 95°C for 10 s, 60°C for 10 s, 72°C for 10 s followed by 10 min of 72°C .

In vitro luciferase assay. The 5' region of *Egfr* corresponding to that described in *A. mellifera*³⁹ was amplified using the 'Luciferase construct PCR primers' (Supplementary Table 1) generating a 378-bp fragment of the 5' region of the *C. floridanus Egfr* gene (scaffold 550: bp113394-bp113772; located using GBrowse of the Hymenoptera Genome Database⁶⁷). *BamHI* and *HindIII* restriction sites were incorporated into the primers in order to generate restriction sites to clone the fragment into the CpG-less pCpG (ref. 70) in 5' to 3' (sense) or 3' to 5' (antisense) orientation, respectively. As the vector does not contain CpG sites, all methylated sites are contained in the *Egfr* 5' region. The constructs were methylated *in vitro* with *SssI* CpG DNA methyltransferases (M0226L, New England Biolabs) as previously described⁶⁸. Transfections were performed using calcium phosphate precipitation as described previously into HEK293 cells (CRL-1573, American Type Culture Collection). Cells were harvested 48 h after transfection and luciferase activity was assayed using the Luciferase Assay System (Promega).

Western blotting analyses. Fifty larvae treated with dimethylsulphoxide or EGFR inhibitor were collected 24 h after treatment and homogenized in RIPA buffer (1 \times PBS, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS and 1 \times complete protease inhibitors; Roche Diagnostics). Total protein yield was determined using Bradford assay (Bio-Rad) and a total of 30 μ g of protein was loaded onto a 10% SDS-polyacrylamide gel electrophoresis. As adapted from bees³⁹, protein extracts were immunoblotted with anti-EGFR (sc-33746, Santa Cruz Biotech) and anti-phosphotyrosine PY20 antibody (525295, Millipore) at 1:1,000 dilution. EGFR was then blotted by a secondary anti-rabbit (sc-2004, Santa Cruz Biotech) and PY20 was blotted by anti-mouse at 1:5,000 dilution. For quantification, total EGFR was determined with anti-EGFR followed by stripping at pH 6.8 and re-blotting with anti-phosphotyrosine in triplicates to determine overall tyrosine phosphorylation. The intensities of the signals were then quantified using Image J software and total phosphorylation was measured and presented as the ratio between the intensity of the phosphorylated band corresponding to EGFR divided by intensity of total EGFR.

EGFR genotyping in adults. Minor ($n = 9$) and major ($n = 8$) representing the extremes of adult development (a and e, see Fig. 1B). Abdomens were removed from adults to exclude exogenous DNA from gastrointestinal microbiota. Template DNA was used with primers used to amplify the EGFR locus used for luciferase promoter constructs (see Supplementary Table 1). Cycling conditions involved an initial step of 5 min at 95°C followed by 35 cycles of 95°C for 1 min, 60°C for 2.5 min, 72°C 1 min, followed by 5 min of 72°C . PCR products were performed at the G enome Qu ebec Innovation Centre (Montr eal, Canada) using both forward and reverse primers.

Statistical analysis. All data are expressed as mean \pm s.e.m., except developmental time, which is expressed as standard deviation (s.d.) and morphometrics in the form of box-and-whisker plots. Comparisons between groups were performed using two-tailed, unpaired Student's *t*-test, except in the case of unequal variance for which the Mann-Whitney *U*-test was performed. Multivariate analysis of variance performed considered all morphometric traits, significance was set at $P < 0.05$ and Pillai's Trace was used. Correlations of percent methylation with terminal larval size were calculated using Pearson correlation and significance was initially set at $P < 0.05$ followed by Bonferroni correction for multiple tests, implemented to correct for multiple comparisons. Statistical analysis was undertaken using Prism (GraphPad Software Inc.) except for multivariate analysis of variance, which was performed in R.

References

- Falconer, D. S. & Mackay, T. F. C. *Introduction to Quantitative Genetics*, 4 edn (Longmans Green, 1996).
- Lynch, M. & Walsh, B. *Genetics and Analysis of Quantitative Traits* (Sinauer Associates, 1998).
- Roff, D. A. A centennial celebration for quantitative genetics. *Evolution* **61**, 1017–1032 (2007).
- Flint, J. & Mackay, T. F. C. Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Res.* **19**, 723–733 (2009).
- Eichler, E. E. *et al.* Missing heritability and strategies for finding the underlying causes of complex disease. *Nat. Rev. Genet.* **11**, 446–450 (2010).
- Schmitz, R. J., Sung, S. & Amasino, R. M. Histone arginine methylation is required for vernalization-induced epigenetic silencing of FLC in winter-annual *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA* **105**, 411–416 (2008).
- Kucharski, R., Maleszka, J., Foret, S. & Maleszka, R. Nutritional control of reproductive status in honeybees via DNA methylation. *Science* **319**, 1827–1830 (2008).
- Weaver, I. C. *et al.* Epigenetic programming by maternal behavior. *Nat. Neurosci.* **7**, 847–854 (2004).
- Jaenisch, R. & Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature Genet.* **33**, 245–254 (2003).
- Herb, B. R. *et al.* Reversible switching between epigenetic states in honeybee behavioral subcastes. *Nat. Neurosci.* **10** (2012).
- Wilson, E. O. The origin and evolution of polymorphism in ants. *Q. Rev. Biol.* **28**, 136–156 (1953).
- Hölldobler, B. & Wilson, E. O. *The Ants* (Belknap Press of Harvard Univ., 1990).
- Wilson, E. O. Caste and division of labor in leaf-cutter ants (Hymenoptera, Formicidae, Atta). I. The overall pattern in *Atta Sexdens*. *Behav. Ecol. Sociobiol.* **7**, 143–156 (1980).
- Klowden, M. J. *Physiological Systems in Insects*, 3rd edn (Academic Press, 2013).
- Nijhout, H. F. The control of body size in insects. *Dev. Biol.* **261**, 1–9 (2003).
- Rajakumar, R. *et al.* Ancestral developmental potential facilitates parallel evolution in ants. *Science* **335**, 79–82 (2012).
- Wheeler, D. E. & Nijhout, H. F. Soldier determination in *Pheidole bicarinata*—inhibition by adult soldiers. *J. Insect Physiol.* **30**, 127–135 (1984).
- Wheeler, W. M. A neglected factor in evolution. *Science* **15**, 766–774 (1902).
- Goetsch, W. Die entstehung der 'soldaten' im ameisenstaat. *Naturwissenschaften* **25**, 803–808 (1937).
- Hughes, W. O., Sumner, S., Van Borm, S. & Boomsma, J. J. Worker caste polymorphism has a genetic basis in *Acromyrmex* leaf-cutting ants. *Proc. Natl Acad. Sci. USA* **100**, 9394–9397 (2003).
- Smith, C. R., Anderson, K. E., Tillberg, C. V., Gadau, J. & Suarez, A. V. Caste determination in a polymorphic social insect: nutritional, social, and genetic factors. *Am. Nat.* **172**, 497–507 (2008).
- Wiernasz, D. C. & Cole, B. J. Patriline shifting leads to apparent genetic caste determination in harvester ants. *Proc. Natl Acad. Sci. USA* **107**, 12958–12962 (2010).
- Smith, F. Effect of reduced food supply upon the stature of *Camponotus* ants. *Entomol. News* **53**, 133–135 (1942).
- Smith, F. Nutritional requirements of *Camponotus* ants. *Ann. Entomol. Soc. Am.* **37**, 401–408 (1944).
- Gadau, J., Heinze, J., Hölldobler, B. & Schmid, M. Population and colony structure of the carpenter ant *Camponotus floridanus*. *Mol. Ecol.* **5**, 785–792 (1996).
- Bonasio, R. *et al.* Genome-wide and caste-specific DNA methylomes of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Curr. Biol.* **22**, 1755–1764 (2012).
- Bonasio, R. *et al.* Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Science* **329**, 1068–1071 (2010).
- Zemach, A., McDaniel, I. E., Silva, P. & Zilberman, D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **328**, 916–919 (2010).
- Feng, S. *et al.* Conservation and divergence of methylation patterning in plants and animals. *Proc. Natl Acad. Sci. USA* **107**, 8689–8694 (2010).
- Razin, A. & Riggs, A. D. DNA methylation and gene function. *Science* **210**, 604–610 (1980).
- Hunt, B. G., Glastad, K. M., Yi, S. V. & Goodisman, M. A. The function of intragenic DNA methylation: insights from insect epigenomes. *Integr. Comp. Biol.* **53**, 319–328 (2013).
- Lyko, F. *et al.* The honey bee epigenomes: differential methylation of brain DNA in queens and workers. *PLoS Biol.* **8**, e1000506 (2010).
- Foret, S. *et al.* DNA methylation dynamics, metabolic fluxes, gene splicing, and alternative phenotypes in honey bees. *Proc. Natl Acad. Sci. USA* **109**, 4968–4973 (2012).
- Li-Byarlay, H. *et al.* RNA interference knockdown of DNA methyl-transferase 3 affects gene alternative splicing in the honey bee. *Proc. Natl Acad. Sci. USA* **110**, 12750–12755 (2013).
- Patel, A. *et al.* The making of a queen: TOR pathway is a key player in diphenic caste development. *PLoS ONE* **2**, e509 (2007).
- Nan, X. *et al.* Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**, 386–389 (1998).
- Juttermann, R., Li, E. & Jaenisch, R. Toxicity of 5-aza-2'-deoxycytidine to mammalian cells is mediated primarily by covalent trapping of DNA methyltransferase rather than DNA demethylation. *Proc. Natl Acad. Sci. USA* **91**, 11797–11801 (1994).
- Smith, S. S., Kaplan, B. E., Sowers, L. C. & Newman, E. M. Mechanism of human methyl-directed DNA methyltransferase and the fidelity of cytosine methylation. *Proc. Natl Acad. Sci. USA* **89**, 4744–4748 (1992).
- Kamakura, M. Royalactin induces queen differentiation in honeybees. *Nature* **473**, 478–483 (2011).
- Mutti, N. S., Wang, Y., Kaftanoglu, O. & Amdam, G. V. Honey bee PTEN—description, developmental knockdown, and tissue-specific expression of splice-variants correlated with alternative social phenotypes. *PLoS ONE* **6**, e22195 (2011).
- Mutti, N. S. *et al.* IRS and TOR nutrient-signaling pathways act via juvenile hormone to influence honey bee caste fate. *J. Exp. Biol.* **214**, 3977–3984 (2011).
- Jorissen, R. N. *et al.* Epidermal growth factor receptor: mechanisms of activation and signalling. *Exp. Cell Res.* **284**, 31–53 (2003).
- Turner, T. L., Stewart, A. D., Fields, A. T., Rice, W. R. & Tarone, A. M. Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in *Drosophila melanogaster*. *PLoS Genet.* **7**, e1001336 (2011).
- Fry, D. W. *et al.* A specific inhibitor of the epidermal growth factor receptor tyrosine kinase. *Science* **265**, 1093–1095 (1994).
- Lubin, F. D., Roth, T. L. & Sweatt, J. D. Epigenetic regulation of *bdnf* gene transcription in the consolidation of fear memory. *J. Neurosci.* **28**, 10576–10586 (2008).
- Brodsky, L., Lee, Y. W. & Costa, M. 5-Azacytidine induces transgene silencing by DNA methylation in Chinese hamster cells. *Mol. Cell Biol.* **19**, 3198–3204 (1999).
- Wheeler, D. E. & Nijhout, H. F. Soldier determination in ants: new role for juvenile hormone. *Science* **213**, 361–363 (1981).
- Simola, D. F. *et al.* A chromatin link to caste identity in the carpenter ant *Camponotus floridanus*. *Genome Res.* **23**, 486–496 (2013).
- Samudio-Ruiz, S. L. & Hudson, L. G. Increased DNA methyltransferase activity and DNA methylation following epidermal growth factor stimulation in ovarian cancer cells. *Epigenetics* **7**, 216–224 (2012).
- Bonduriansky, R., Crean, A. J. & Day, T. The implications of nongenetic inheritance for evolution in changing environments. *Evol. Appl.* **5**, 192–201 (2012).
- Johannes, F. *et al.* Assessing the impact of transgenerational epigenetic variation on complex traits. *PLoS Genet.* **5**, e1000530 (2009).
- Danchin, E. *et al.* Beyond DNA: integrating inclusive inheritance into an extended theory of evolution. *Nat. Rev. Genet.* **12**, 475–486 (2011).
- Waterland, R. A. & Jirtle, R. L. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol. Cell Biol.* **23**, 5293–5300 (2003).
- Toperoff, G. *et al.* Genome-wide survey reveals predisposing diabetes type 2-related DNA methylation variations in human peripheral blood. *Hum. Mol. Genet.* **21**, 371–383 (2012).
- Provine, W. B. *The Origins of Theoretical Population Genetics*, 2nd edn (Univ. Chicago, 2001).
- Rockman, M. V. The Qtn program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* **66**, 1–17 (2012).
- Farrall, M. Quantitative genetic variation: a post-modern view. *Hum. Mol. Genet.* **13**, R1–R7 (2004).
- Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010).
- Kroymann, J. & Mitchell-Olds, T. Epistasis and balanced polymorphism influencing complex trait variation. *Nature* **435**, 95–98 (2005).

60. Steinmetz, L. M. *et al.* Dissecting the architecture of a quantitative trait locus in yeast. *Nature* **416**, 326–330 (2002).
61. Dartigues, D. & Passera, L. Larval polymorphism and the apparition of female castes in *Camponotus-Aethiops Latreille* (Hymenoptera, Formicidae). *Bulletin De La Societe Zoologique De France* **104**, 197–207 (1979).
62. Solis, D. R. *et al.* Morphological description of the immatures of the ant, *Monomorium floricola*. *J. Insect Sci.* **10**, 15 (2010).
63. Wheeler, G. C. & Wheeler, J. Ant larvae: review and synthesis. *Entomol. Soc. Washington* **7**, 1–108 (1976).
64. Diniz-Filho, J. A. F., Vonzuben, C. J., Fowler, H. G., Schindwein, M. N. & Bueno, O. C. Multivariate morphometrics and allometry in a polymorphic ant. *Insectes Soc.* **41**, 153–163 (1994).
65. Fraser, V. S., Kaufmann, B., Oldroyd, B. P. & Crozier, R. H. Genetic influence on caste in the ant *Camponotus consobrinus*. *Behav. Ecol. Sociobiol.* **47**, 188–194 (2000).
66. Karimi, M. *et al.* LUMA (LUMinometric Methylation Assay)--a high throughput method to the analysis of genomic DNA methylation. *Exp. Cell Res.* **312**, 1989–1995 (2006).
67. Munoz-Torres, M. C. *et al.* Hymenoptera Genome Database: integrated community resources for insect species of the order Hymenoptera. *Nucleic Acids Res.* **39**, D658–D662 (2011).
68. Alvarado, S., Wyglinski, J., Suderman, M., Andrews, S. A. & Szyf, M. Methylated DNA binding domain protein 2 (MBD2) coordinately silences gene expression through activation of the microRNA hsa-mir-496 promoter in breast cancer cell line. *PLoS ONE* **8**, e74009 (2013).
69. Lourenço, A. P., Mackert, A., dos Santos Cristino, A. & Simões, Z. L. P. Validation of reference genes for gene expression studies in the honey bee, *Apis mellifera*, by quantitative real-time RT-PCR. *Adipologie* **39**, 372–385 (2008).
70. Klug, M. & Rehli, M. Functional analysis of promoter CpG methylation using a CpG-free luciferase reporter vector. *Epigenetics* **1**, 127–130 (2006).

Acknowledgements

We thank M. Tajerian for inspiring this collaboration. We also thank D. Roff, R. Barrett, Y. Idaghdour, E. Despland and Abouheif Lab members for comments on the manuscript. Finally, we thank W. Tschinkel, J. King and L. Davis for help with collection and identification and S. Silvestrin, E. Lo, B. Fung, T. Chen, Y. Chen and J. Fan for help with colonies. Thanks to Dominic Ouellette and Melanie Couture for taking photos for panel 1B. This work was supported by the Sackler McGill programme in psychobiology and epigenetics to M.S. a grant from the Canadian Institute of Health Research to M.S. (MOP-42411) and Canada Research Chairs programme and NSERC Discovery grants to E.A.

Author contributions

S.A. and R.R. contributed equally to this work. S.A., R.R., E.A. and M.S. conceived and designed the experiments. S.A. and R.R. performed the experiments. S.A., R.R., E.A. and M.S. analysed the data. S.A., R.R., E.A., and M.S. wrote the paper.

Additional information

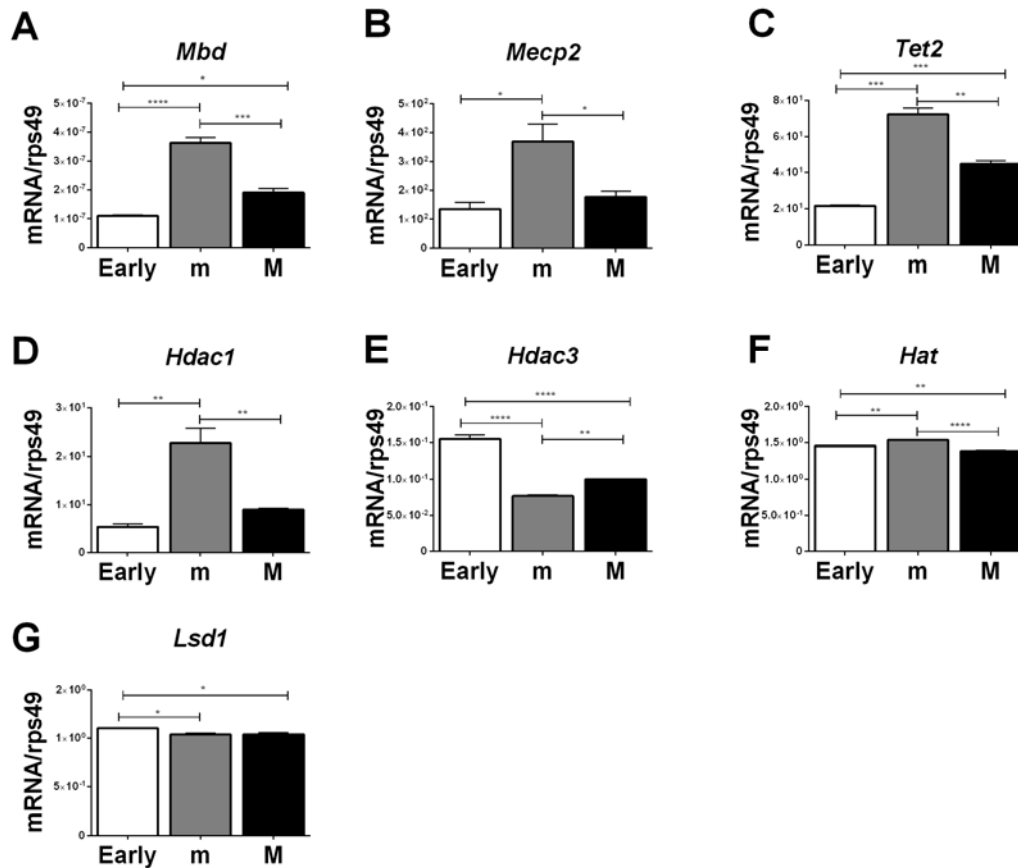
Accession codes. The *Egfr* promoter sequence generated in this study has been deposited in NCBI nucleotide database under accession code KP325210.

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

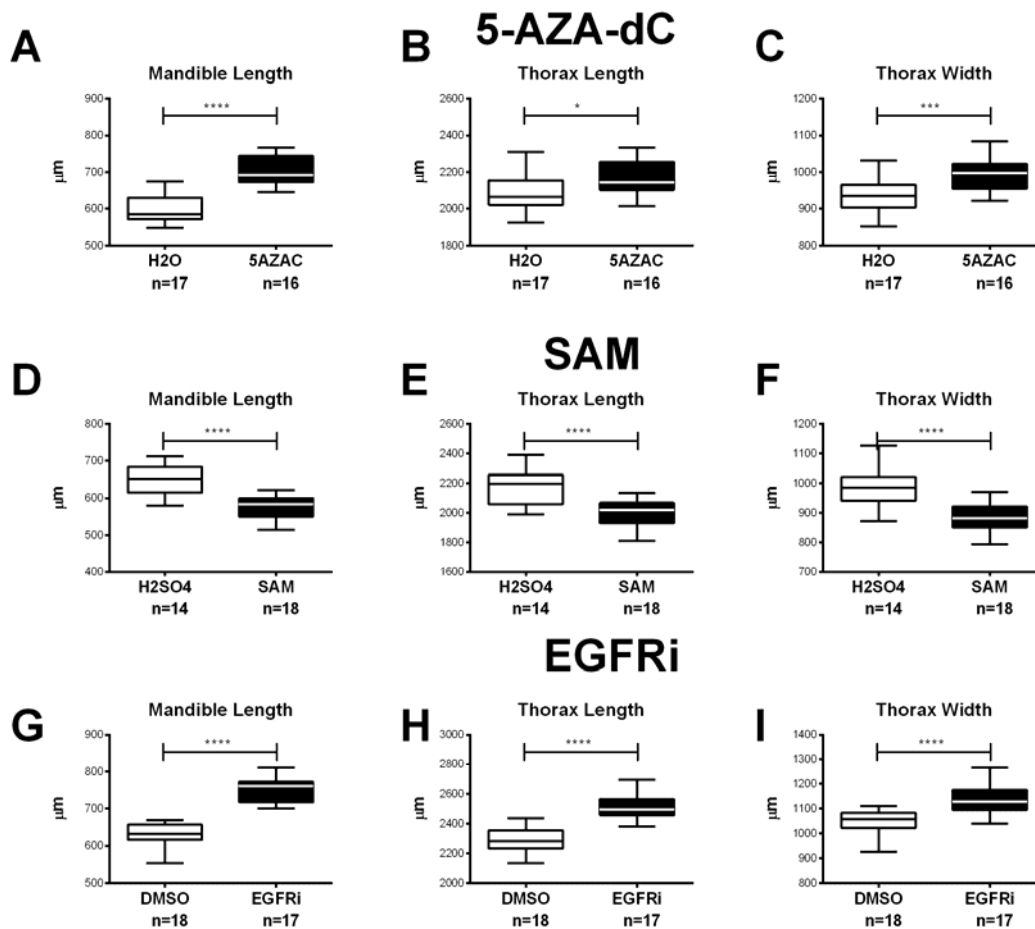
Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

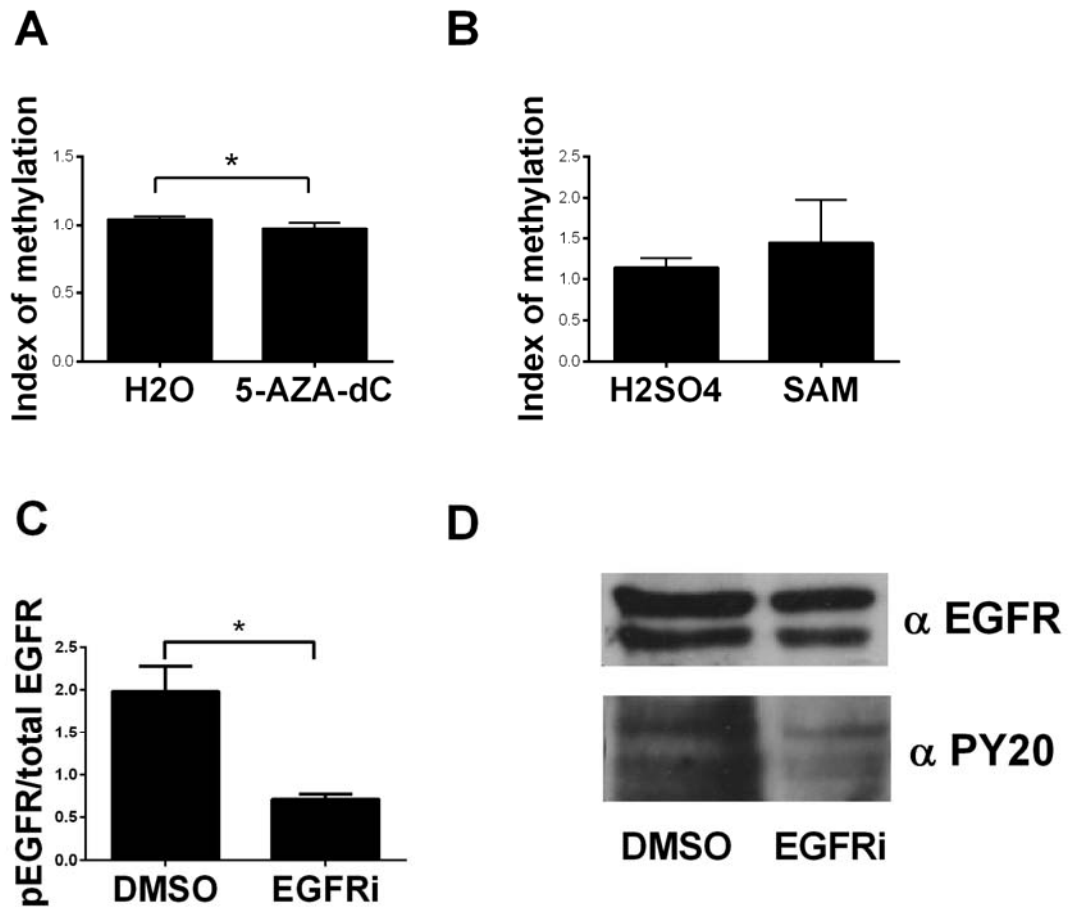
How to cite this article: Alvarado, S. *et al.* Epigenetic variation in the *Egfr* gene generates quantitative variation in a complex trait in ants. *Nat. Commun.* **6**:6513 doi: 10.1038/ncomms7513 (2015).



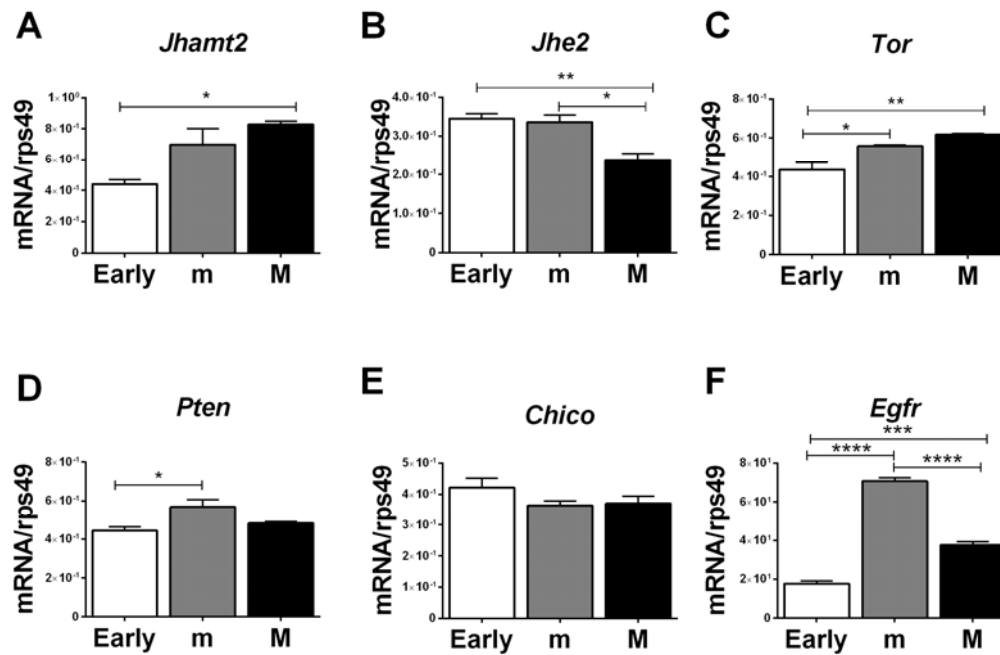
Supplementary Figure 1: Quantitative gene expression of additional DNA methylation and histone modification regulators. Expression of (A) *Mbd*, (B) *Mecp2*, (C) *Tet2* (D) *Hdac1*, (E) *Hdac3*, (F) *Hat*, (G) *Lsd1* in early 4th instar larvae (white bar), late 4th instar minor worker larvae (grey bar) and late 4th instar major worker larvae (black bar). Bars indicate mean, while error bars indicate \pm s.e.m. Statistical significance values for Student's t-test are as follows: *= p<0.05, **=p<0.01, ***=p<0.001, ****=p<0.0001. Sample sizes for each larval category is n \geq 30 individual larvae.



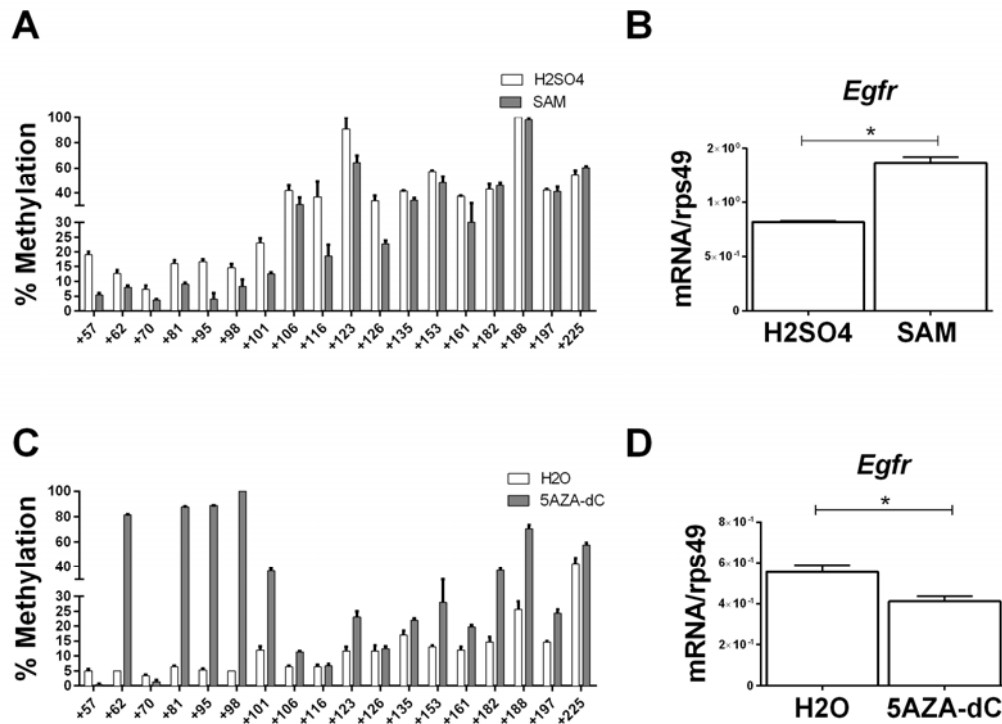
Supplementary Figure 2: Additional morphometric measurements of 5-AZA-dC, SAM, and EGFRi treatments. (A) Mandible length (B) Thorax length and (C) Thorax width of 5-AZA-dC treated animals (black box)(n=17) and controls (white box; H2O)(n=16). (D) Mandible length, (E) Thorax length and (F) Thorax width in SAM treated animals (black box)(n=18) and controls (white box; H2SO4)(n=14). (G) Mandible length, (H) Thorax length and (I) Thorax width in EGFR inhibited animals (black box)(n=17) and controls (white box; DMSO)(n=18). Boxplot whiskers indicate min and max. Box defined by 25th percentile, mode, and 75th percentile. Statistical significance values for Student's t-test are as follows: *= p<0.05, ***=p<0.001, ****=p<0.0001. All measurement data are normally distributed (tested with Shapiro-Wilk's test). To consider all 5 measurements together, we used a MANOVA analysis to statistically test for differences of treatments with controls across the animal (including scape length and head width of Fig.2 and Fig. 3). MANOVA statistics are as follows: SAM (pillai trace=0.76571, F= 17.649, p=3.045e-09); 5-AZA-dC (pillai trace=0.81913, F= 24.456, p=3.045e-09); EGFRi (pillai trace= 0.85263, F= 33.556, p=3.293e-11).



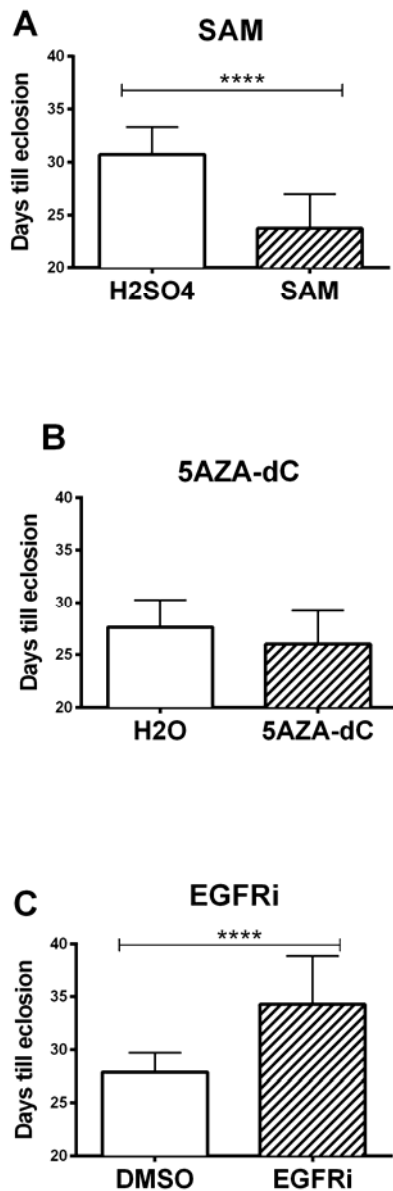
Supplementary Figure 3: Validation of drug treatments for 5-AZA-dC, SAM and EGFRi treatments. (A) Genomic methylation of 5-AZA-dC treated early 4th instar larvae represented as an index of [MspI/HpaII] restriction enzyme cutting quantified by LUMA (n=50). (B) Genomic methylation of SAM treated early 4th instar larvae represented as an index of [MspI/HpaII] restriction enzyme cutting quantified by LUMA(n=50). (C) Dephosphorylation of EGFR following 24 hours inhibition of early 4th instar larvae (n=50), quantified from Western blot. (D) Qualitative Western blot analysis of DMSO and EGFRi treated animals. Bars indicate mean, while error bars indicate \pm s.e.m. Statistical significance values for Student's t-test are as follows: *= p<0.05.



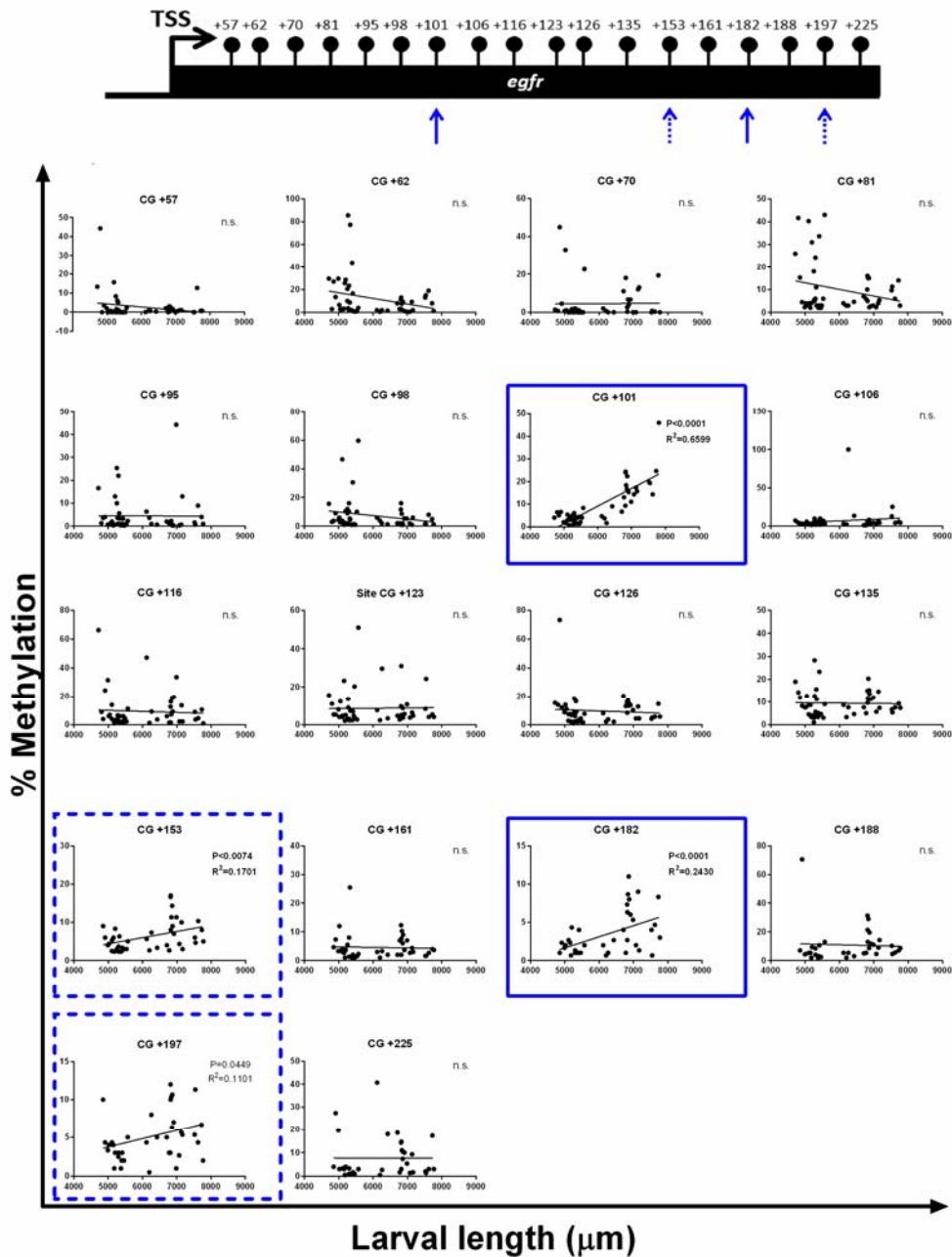
Supplementary Figure 4: Quantitative gene expression of key regulators of growth. (A) *JHAMT2*, (B) *JHE2*, (C) *Tor*, (D) *Pten*, (E) *chico*, (F) *Egfr* in early 4th instar larvae (white bar), late 4th instar minor worker larvae (grey bar) and late 4th instar major worker larvae (black bar). Bars indicate mean, while error bars indicate \pm s.e.m. Statistical significance values for Student's t-test are as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$. Sample sizes for each larval category is pooled from $n \geq 30$ individual larvae.



Supplementary Figure 5: Methylation and quantitative expression of *Egfr* following SAM and 5-AZA-dC administration. (A) Bisulfite map showing percent (%) methylation of *Egfr* in control (H2SO4; white bars) and SAM-treated animals (grey bars) for each of the 18 sites within the first 225bp and (B) expression of *Egfr* following SAM administration relative to control (H2SO4). (C) Bisulfite map showing percent (%) methylation for control (H2O; white bars) and 5-AZA-dC (grey bars) for each of the 18 sites within the first 225bp (D) expression of *Egfr* following 5-AZA-dC administration relative to control (H2O). Bars indicate mean, while error bars indicate \pm s.e.m. Statistical significance values for Student's t-test are as follows: * = $p < 0.05$. Sample sizes for each larval category is pooled from $n \geq 20$ individual larvae.



Supplementary Figure 6: Developmental time is affected by global (genome-wide) DNA methylation and EGFR signaling. Days till eclosion (time from treatment of early 4th instar larvae till the beginning of adulthood) for worker larvae after (A) SAM administration (hatched bar)(n=14) relative to control (H₂SO₄; white bar)(n=18), (B) 5-AZA-dC administration (hatched bar)(n=16) relative to control (H₂O; white bar)(n=17), and (C) EGFR inhibition (hatched bar)(n=17) relative to control (DMSO; white bar)(n=18). Frequency of eclosion timing was not always normally distributed (tested with Shapiro-Wilk's test). Therefore the Mann-Whitney U test was used to compare controls with respective pharmacological manipulation. Bars indicate mean and error bars indicate standard deviation. Statistical significance values for Student's t-test are as follows: ****=p<0.0001.



Supplementary Figure. 7: Association of percent (%) methylation of *Egfr* CpG sites and final size of workers. Map of CpG dinucleotides at top (methylated sites indicated by black lollipops). Regression of percent (%) methylation (y-axis) versus terminal larval length in μm (x-axis) for all 18 CpG dinucleotides occurring in the first 225bp of *Egfr* (n=50). Regressions of CpG sites +101, +153, +182 and +197 with larval size were significant (Table S2). Following Bonferroni correction for multiple comparisons, sites +101 and +182 remain significant (solid blue box and arrows), while site +153 and +197 became marginally insignificant (dashed box and arrow). Refer to Table S2 for all statistical information.


```

minor_1      TCTGCCATATCAAGACGATCAATTGGGAGGAAATAATTACCGGTCCGGGAGGCCGGTACT 60
minor_2      TCTGCCATATCAAGACGATCAATTGGGAGGAAATAATTACCGGTCCGGGAGGCCGGTACT 60
minor_3      TCTGCCATATCAAGACGATCAATTGGGAGGAAATAATTACCGGTCCGGGAGGCCGGTACT 60
minor_4      TCTGCCATATCAAGACGATCAATTGGGAGGAAATAATTACCGGTCCGGGAGGCCGGTACT 60
minor_5      TCTGCCATATCAAGACGATCAATTGGGAGGAAATAATTACCGGTCCGGGAGGCCGGTACT 60
minor_6      TCTGCCATATCAAGACGATCAATTGGGAGGAAATAATTACCGGTCCGGGAGGCCGGTACT 60
minor_7      TCTGCCATATCAAGACGATCAATTGGGAGGAAATAATTACCGGTCCGGGAGGCCGGTACT 60
minor_8      TCTGCCATATCAAGACGATCAATTGGGAGGAAATAATTACCGGTCCGGGAGGCCGGTACT 60
minor_9      TCTGCCATATCAAGACGATCAATTGGGAGGAAATAATTACCGGTCCGGGAGGCCGGTACT 60
ma_jor_10   TCTGCCATATCAAGACGATCAATTGGGAGGAAATAATTACCGGTCCGGGAGGCCGGTACT 60
ma_jor_11   TCTGCCATATCAAGACGATCAATTGGGAGGAAATAATTACCGGTCCGGGAGGCCGGTACT 60
ma_jor_12   TCTGCCATATCAAGACGATCAATTGGGAGGAAATAATTACCGGTCCGGGAGGCCGGTACT 60
ma_jor_13   TCTGCCATATCAAGACGATCAATTGGGAGGAAATAATTACCGGTCCGGGAGGCCGGTACT 60
ma_jor_14   TCTGCCATATCAAGACGATCAATTGGGAGGAAATAATTACCGGTCCGGGAGGCCGGTACT 60
ma_jor_15   TCTGCCATATCAAGACGATCAATTGGGAGGAAATAATTACCGGTCCGGGAGGCCGGTACT 60
ma_jor_16   TCTGCCATATCAAGACGATCAATTGGGAGGAAATAATTACCGGTCCGGGAGGCCGGTACT 60
ma_jor_17   TCTGCCATATCAAGACGATCAATTGGGAGGAAATAATTACCGGTCCGGGAGGCCGGTACT 60
*****
                                  101
minor_1      TTTACGTTGACAATTTTACGTCGCCGGAAACGCAATTGTCCGGAATGCGACGAGAGCTGCG 120
minor_2      TTTACGTTGACAATTTTACGTCGCCGGAAACGCAATTGTCCGGAATGCGACGAGAGCTGCG 120
minor_3      TTTACGTTGACAATTTTACGTCGCCGGAAACGCAATTGTCCGGAATGCGACGAGAGCTGCG 120
minor_4      TTTACGTTGACAATTTTACGTCGCCGGAAACGCAATTGTCCGGAATGCGACGAGAGCTGCG 120
minor_5      TTTACGTTGACAATTTTACGTCGCCGGAAACGCAATTGTCCGGAATGCGACGAGAGCTGCG 120
minor_6      TTTACGTTGACAATTTTACGTCGCCGGAAACGCAATTGTCCGGAATGCGACGAGAGCTGCG 120
minor_7      TTTACGTTGACAATTTTACGTCGCCGGAAACGCAATTGTCCGGAATGCGACGAGAGCTGCG 120
minor_8      TTTACGTTGACAATTTTACGTCGCCGGAAACGCAATTGTCCGGAATGCGACGAGAGCTGCG 120
minor_9      TTTACGTTGACAATTTTACGTCGCCGGAAACGCAATTGTCCGGAATGCGACGAGAGCTGCG 120
ma_jor_10   TTTACGTTGACAATTTTACGTCGCCGGAAACGCAATTGTCCGGAATGCGACGAGAGCTGCG 120
ma_jor_11   TTTACGTTGACAATTTTACGTCGCCGGAAACGCAATTGTCCGGAATGCGACGAGAGCTGCG 120
ma_jor_12   TTTACGTTGACAATTTTACGTCGCCGGAAACGCAATTGTCCGGAATGCGACGAGAGCTGCG 120
ma_jor_13   TTTACGTTGACAATTTTACGTCGCCGGAAACGCAATTGTCCGGAATGCGACGAGAGCTGCG 120
ma_jor_14   TTTACGTTGACAATTTTACGTCGCCGGAAACGCAATTGTCCGGAATGCGACGAGAGCTGCG 120
ma_jor_15   TTTACGTTGACAATTTTACGTCGCCGGAAACGCAATTGTCCGGAATGCGACGAGAGCTGCG 120
ma_jor_16   TTTACGTTGACAATTTTACGTCGCCGGAAACGCAATTGTCCGGAATGCGACGAGAGCTGCG 120
ma_jor_17   TTTACGTTGACAATTTTACGTCGCCGGAAACGCAATTGTCCGGAATGCGACGAGAGCTGCG 120
*****
                                  153                                  182
minor_1      AACAGGGTTGCTGGGGCGAGGGTCCGGGAGAAGTGTCAAAAAGTACTCGAAGACGAACTGCT 180
minor_2      AACAGGGTTGCTGGGGCGAGGGTCCGGGAGAAGTGTCAAAAAGTACTCGAAGACGAACTGCT 180
minor_3      AACAGGGTTGCTGGGGCGAGGGTCCGGGAGAAGTGTCAAAAAGTACTCGAAGACGAACTGCT 180
minor_4      AACAGGGTTGCTGGGGCGAGGGTCCGGGAGAAGTGTCAAAAAGTACTCGAAGACGAACTGCT 180
minor_5      AACAGGGTTGCTGGGGCGAGGGTCCGGGAGAAGTGTCAAAAAGTACTCGAAGACGAACTGCT 180
minor_6      AACAGGGTTGCTGGGGCGAGGGTCCGGGAGAAGTGTCAAAAAGTACTCGAAGACGAACTGCT 180
minor_7      AACAGGGTTGCTGGGGCGAGGGTCCGGGAGAAGTGTCAAAAAGTACTCGAAGACGAACTGCT 180
minor_8      AACAGGGTTGCTGGGGCGAGGGTCCGGGAGAAGTGTCAAAAAGTACTCGAAGACGAACTGCT 180
minor_9      AACAGGGTTGCTGGGGCGAGGGTCCGGGAGAAGTGTCAAAAAGTACTCGAAGACGAACTGCT 180
ma_jor_10   AACAGGGTTGCTGGGGCGAGGGTCCGGGAGAAGTGTCAAAAAGTACTCGAAGACGAACTGCT 180
ma_jor_11   AACAGGGTTGCTGGGGCGAGGGTCCGGGAGAAGTGTCAAAAAGTACTCGAAGACGAACTGCT 180
ma_jor_12   AACAGGGTTGCTGGGGCGAGGGTCCGGGAGAAGTGTCAAAAAGTACTCGAAGACGAACTGCT 180
ma_jor_13   AACAGGGTTGCTGGGGCGAGGGTCCGGGAGAAGTGTCAAAAAGTACTCGAAGACGAACTGCT 180
ma_jor_14   AACAGGGTTGCTGGGGCGAGGGTCCGGGAGAAGTGTCAAAAAGTACTCGAAGACGAACTGCT 180
ma_jor_15   AACAGGGTTGCTGGGGCGAGGGTCCGGGAGAAGTGTCAAAAAGTACTCGAAGACGAACTGCT 180
ma_jor_16   AACAGGGTTGCTGGGGCGAGGGTCCGGGAGAAGTGTCAAAAAGTACTCGAAGACGAACTGCT 180
ma_jor_17   AACAGGGTTGCTGGGGCGAGGGTCCGGGAGAAGTGTCAAAAAGTACTCGAAGACGAACTGCT 180
*****
                                  197
minor_1      CGCCTCAGTGTGCGCAGGGCAGGTGTTTCGGTCCCTAATCCACGCGAGTGTGCCATCTTT 240
minor_2      CGCCTCAGTGTGCGCAGGGCAGGTGTTTCGGTCCCTAATCCACGCGAGTGTGCCATCTTT 240
minor_3      CGCCTCAGTGTGCGCAGGGCAGGTGTTTCGGTCCCTAATCCACGCGAGTGTGCCATCTTT 240
minor_4      CGCCTCAGTGTGCGCAGGGCAGGTGTTTCGGTCCCTAATCCACGCGAGTGTGCCATCTTT 240
minor_5      CGCCTCAGTGTGCGCAGGGCAGGTGTTTCGGTCCCTAATCCACGCGAGTGTGCCATCTTT 240
minor_6      CGCCTCAGTGTGCGCAGGGCAGGTGTTTCGGTCCCTAATCCACGCGAGTGTGCCATCTTT 240
minor_7      CGCCTCAGTGTGCGCAGGGCAGGTGTTTCGGTCCCTAATCCACGCGAGTGTGCCATCTTT 240
minor_8      CGCCTCAGTGTGCGCAGGGCAGGTGTTTCGGTCCCTAATCCACGCGAGTGTGCCATCTTT 240
minor_9      CGCCTCAGTGTGCGCAGGGCAGGTGTTTCGGTCCCTAATCCACGCGAGTGTGCCATCTTT 240
ma_jor_10   CGCCTCAGTGTGCGCAGGGCAGGTGTTTCGGTCCCTAATCCACGCGAGTGTGCCATCTTT 240
ma_jor_11   CGCCTCAGTGTGCGCAGGGCAGGTGTTTCGGTCCCTAATCCACGCGAGTGTGCCATCTTT 240
ma_jor_12   CGCCTCAGTGTGCGCAGGGCAGGTGTTTCGGTCCCTAATCCACGCGAGTGTGCCATCTTT 240
ma_jor_13   CGCCTCAGTGTGCGCAGGGCAGGTGTTTCGGTCCCTAATCCACGCGAGTGTGCCATCTTT 240
ma_jor_14   CGCCTCAGTGTGCGCAGGGCAGGTGTTTCGGTCCCTAATCCACGCGAGTGTGCCATCTTT 240
ma_jor_15   CGCCTCAGTGTGCGCAGGGCAGGTGTTTCGGTCCCTAATCCACGCGAGTGTGCCATCTTT 240
ma_jor_16   CGCCTCAGTGTGCGCAGGGCAGGTGTTTCGGTCCCTAATCCACGCGAGTGTGCCATCTTT 240
ma_jor_17   CGCCTCAGTGTGCGCAGGGCAGGTGTTTCGGTCCCTAATCCACGCGAGTGTGCCATCTTT 240
*****

```

Supplementary Figure 8: There is no allelic bias between minor and major workers. Sequence alignment of the EGFR locus studied in minor (n=9) vs major (n=8) workers. Minor and majors were selected from the extremes of sizing phenotypes (a and e)

described in Fig. 1A. Alignment was performed using ClustalW2 multiple sequence alignment. All CpGs are bolded with those assayed for methylation in blue. CpGs correlated with size are highlighted in yellow and numbered to correspond with the physical map of EGFR locus studied.

Supplementary Table 1: Primers used for all experiments

Bisulfite primers	
r1 bis EGFR for	TTTTAGTGTGGTAGGGTAGGTGTT
r1 bis EGFR rev	AAAAATTCTTACAAACAAAACAATC
r1s1 EGFR	GGTAGGGTAGGTGTTT
r2 bis EGFR for	TGGGTATGTATAATAATTATAATTTTGT
r2 bis EGFR rev	ACACCTACCCTACCAACACTAAAAC
r2s1 EGFR	AATTGGGAGGAAATAATTA
qPCR	
Chico-1 for	AGTCACGGGTGCGACTGT
Chico-1 rev	CTGAGTCCGACGAGCACAT
Chico-2 for	CCTCTGGGTGCGACTGTG
Chico-2 rev	CTGAGTCCGACGAGCACAT
Chico-3 for	AGATCCGTTTTACGGTGTGC
Chico-3 rev	TGTCCCTCGTTGTTGGTACA
JHMT1 for	CTAAAACGAGTCTCACAGATTCC
JHMT1 rev	CTTTTGGCCTCTTTGGTTT
DNMT1 for	CTGTGTGCCTTTGACACTGG
DNMT1 rev	TGGCCCCATATCTTTTGTG
DNMT3a for	GACTGCTGCTTGAAGGAACC
DNMT3a rev	TTTGAATGTAGTCGCGCATC
MBD for	GGAATGGATCTGCCAAAGAA
MBD rev	CCGTTTTCGATCCTGTTTGT
TET1 for	AGATAGTTTGCCCGATGGTG
TET1 rev	TTTCGAGAGCTGTCATTCCA
MAPK for	ACGAGGTTTAAGCATGAAAATATAA
MAPK rev	AACAATATATACGTCTTTCATTTGCTC
WDR5 for	TGGGAATTAAGTTCGGGTAAA
WDR5 rev	TGAGATTGGATTGTGGATTAAAA
TOR for	TGCATTAAAGGTAGCAACGGTA
TOR rev	TATCCGGATCTCCCAACAAG
PTEN for	GGTCAAGCATGTCTGCGTTA
PTEN rev	TTCCGAACCTCGTAAACACC
JHE1 for	TTTATCCCGTAGCCGATTTG
JHE1 rev	GCAATCTCCACCTGCTTCTC
JHE2 for	GCTGAATTCATCGCTGACAA
JHE2 rev	GAAAATGCGGACCAAGAAGA
JHMT2 for	TCCAAATGCAGTAATAATGGGTA
JHMT2 rev	TTCGAATTTGAGCTGTTTCTCA
EGFR for	GCACGTACCAGAGGGATGTT
EGFR rev	AAGCCGTATCCTGTGCACTT
LSD1 for	TCGCCACATTTTCGTAAATCA
LSD1 rev	TCTTTTGAACCGTTTGACC

HAT for	CACAAGGCGATTTGAGGTCT
HAT rev	AAAACCAATCGTCGCGTATC
HDAC3 for	GAGTAAAGTCTGGCGCGAAG
HDAC3 rev	ATCGACTTGGCTGCTTCAGT
HDAC1 for	AATTCCTGAGGATGGTGCTG
HDAC1 rev	TCCTGACCTTTTCCAAACG
MECP2 for	TCATGCATCTCGCTCAAAAC
MECP2 rev	AACGGCACCATCCGTAGTAG
Luciferase construct primers	
pcpgl EGFR for S	ggatccTATA CCTCTGCCATATCAAGACGA
pcpgl EGFR rev S	AAGCTTTATA GCAAAAAGACAGAAGAATCGTG
pcpgl EGFR for AS	AAGCTTTATA CCTCTGCCATATCAAGACGA
pcpgl EGFR rev AS	ggatccTATA GCAAAAAGACAGAAGAATCGTG

Supplementary Table 2: Parameter estimates and Bonferroni-corrected statistical conclusions for regression analyses between percent methylation of 18 CpG sites within first 225bp of Egfr and final larval size (see Supplemental figure 5)

CpG	R-squared	p-value	Bonferroni-corrected statistical conclusions (experiment-wide α value = 0.00278)
+57	0.04525	0.1381	n.s.
+62	0.06923	0.0649	n.s.
+70	0.0001212	0.9395	n.s.
+81	0.05805	0.0919	n.s.
+95	0.0000637	0.9561	n.s.
+98	0.04392	0.1441	n.s.
+101	0.6599	<0.0001	significant
+106	0.01769	0.3572	n.s.
+116	0.002905	0.7107	n.s.
+123	0.0002577	0.9119	n.s.
+126	0.005211	0.6184	n.s.
+135	0.00005444	0.9603	n.s.
+153	0.1701	0.0074	n.s.
+161	0.003666	0.7069	n.s.
+182	0.243	0.0019	significant
+188	0.0002717	0.9229	n.s.
+197	0.1101	0.0449	n.s.
+225	0.001243	0.8359	n.s.