DNA methylation regulates transcriptional homeostasis of algal endosymbiosis in the coral model Aiptasia

Yong Li[1], Yi Jin Liew[1], Guoxin Cui[1], Maha J. Cziesielski[1], Noura Zahran[1], Craig T. Michell[1], Christian R. Voolstra[1], Manuel Aranda[1]

[1]Red Sea Research Center, Division of Biological and Environmental Science and Engineering, King Abdullah University of Science and Technology (KAUST), Thuwal, KSA

*Correspondence to: manuel.aranda@kaust.edu.sa

Running title: DNA methylation in cnidarian-algal symbiosis

Key words: DNA methylation, Aiptasia, cnidarian-dinoflagellate symbiosis, coral bleaching

Abstract

The symbiotic relationship between cnidarians and dinoflagellates is the cornerstone of coral reef ecosystems. Although research is focusing on the molecular mechanisms underlying this symbiosis, the role of epigenetic mechanisms, which have been implicated in transcriptional regulation and acclimation to environmental change, is unknown. To assess the role of DNA methylation in the cnidarian-dinoflagellate symbiosis, we analyzed genome-wide CpG methylation, histone associations, and transcriptomic states of symbiotic and aposymbiotic anemones in the model system Aiptasia. We find methylated genes are marked by histone H3K36me3 and show significant reduction of spurious transcription and transcriptional noise, revealing a role of DNA methylation in the maintenance of transcriptional homeostasis. Changes in DNA methylation and expression show enrichment for symbiosis-related processes such as immunity, apoptosis, phagocytosis recognition and phagosome formation, and unveil intricate interactions between the underlying pathways. Our results demonstrate that DNA methylation provides an epigenetic mechanism of transcriptional homeostasis during symbiosis.

Introduction

Coral reefs are ecologically important marine ecosystems, which cover less than 0.2% of our oceans but sustain an estimated ~25% of the world's marine species and 32 of 33 animal phyla (Spalding and Grenfell 1997; Davidson 2002; Sylvain 2006). Coral reefs are also economically important by providing food and livelihood opportunities to at least 500 million people; worldwide, they have a net present value of almost USD 800 billion, and they generate USD 30 billion in net economic benefits annually (Sylvain 2006). Unfortunately, these

ecosystems are under severe threat from anthropogenic stressors including global warming and water pollution, among others, which can cause coral bleaching (loss of intracellular endosymbionts from coral) and overall coral reef decline. Despite increasing efforts on studying the mechanisms underlying the regulation and environmental stress related breakdown of this symbiotic association (Davy et al. 2012; Meyer and Weis 2012), we still lack knowledge on basic molecular processes, for instance whether epigenetic mechanisms are involved in symbiosis regulation and could potentially contribute to increased resilience in response to environmental stress as reported in other organisms (Rando and Verstrepen 2007; Lämke and Bäurle 2017).

DNA methylation plays an important role in many biological processes of plants and animals (Bird 2002; Suzuki and Bird 2008; He et al. 2011; Jones 2012). It has been proposed as a mechanism for organisms to adjust their phenotype in response to their environment in order to optimize organismal response to changing environmental conditions (Richards 2006; Rando and Verstrepen 2007). For instance, recent findings in mice show an important function for DNA methylation in inhibiting spurious transcription along the gene body, allowing for reduction of nonsense transcripts from highly expressed loci (Neri et al. 2017). Similar functions have also been proposed in plants, suggesting a general role of DNA methylation in the maintenance of transcriptional homeostasis (Zilberman 2017).

Several studies on DNA methylation in cnidarians have been published recently (Dixon et al. 2016; Putnam et al. 2016), however, its role and function in cnidarians is, at present, mainly unknown (Torda et al. 2017). The sea anemone Aiptasia is an emerging model to study the cnidarian-dinoflagellate symbiosis. Like corals, it establishes a stable but temperature sensitive symbiosis with dinoflagellates of the genus *Symbiodinium* but, unlike corals, can also be

naturally maintained in an aposymbiotic state. This, compounded with its ease of culture, provides a tractable system to study the molecular mechanism underlying symbiosis without the impeding stress responses associated with coral bleaching stress (Voolstra 2013; Baumgarten et al. 2015).

Using the model system Aiptasia (strain CC7, sensu *Exaiptasia pallida*), we obtained whole-genome CpG DNA methylation, ChIP-Seq and RNA-Seq data from aposymbiotic (Apo) and symbiotic (Sym) individuals to study the function of DNA methylation in transcriptional regulation and its role in the cnidarian-dinoflagellate symbiosis.

Results

Aiptasia DNA Methylation patterns change with symbiotic states

To assess changes in DNA methylation in response to symbiosis, we performed whole-genome bisulfite sequencing with an average coverage of 53× per individual on 12 anemones, providing 6 biological replicates per treatment (symbiotic vs. aposymbiotic). Methylation calling using the combined dataset identified 710,768 CpGs (6.37% of all CpGs in Aiptasia genome), i.e. methylated sites in the Aiptasia genome. Notably, the percentage of CpGs is much lower than in mammals (60–90%) (Tucker 2001), but comparable to the coral *Stylophora pistillata* (7%) (Liew et al. 2017). We identified 10,822 genes (37% of all 29,269 gene models identified in the Aiptasia genome) with at least 5 methylated positions that were subsequently defined as methylated genes. On average, these genes had 18.4% CpGs methylated, 3-fold higher than the average methylation density across the entire genome (Chi-squared test p value $< 2.2 \times 10^{-16}$) and 167-fold higher than the methylation levels in non-coding regions. These findings indicate

79 that the distribution of CpG methylation is non-random and mainly located in gene bodies,
80 similar to corals (Dixon et al. 2017; Liew et al. 2017) and other invertebrate species (Feng et al.
81 2010; Gavery and Roberts 2010; Wang et al. 2013; Gonzalez-Romero et al. 2017).

82 To analyze the relationship between methylation density (percentage of CpGs) and gene
83 density (the number of genes per 10,000 bp), we ran a sliding window (window size: 40 kb, step:
84 30 kb) and visualized the results in a Circos plot (Fig. S1) (Krzywinski et al. 2009). The
85 correlation of CpG content and distribution of methylation showed a negative correlation
86 (Pearson correlation coefficient: $r = -0.31$, p value $< 2.2 \times 10^{-16}$) suggesting that methylation
87 tends to preferentially occur in CpG-poor regions (Fig. S2). Gene density had a positive
88 correlation with methylation density ($r = 0.21$, p value $< 2.2 \times 10^{-16}$) consistent with the finding
89 that methylation is predominantly located in gene bodies (Fig. 1). We also observed that within
90 gene bodies, introns showed significantly higher methylation densities than exons (Fig. 1B).

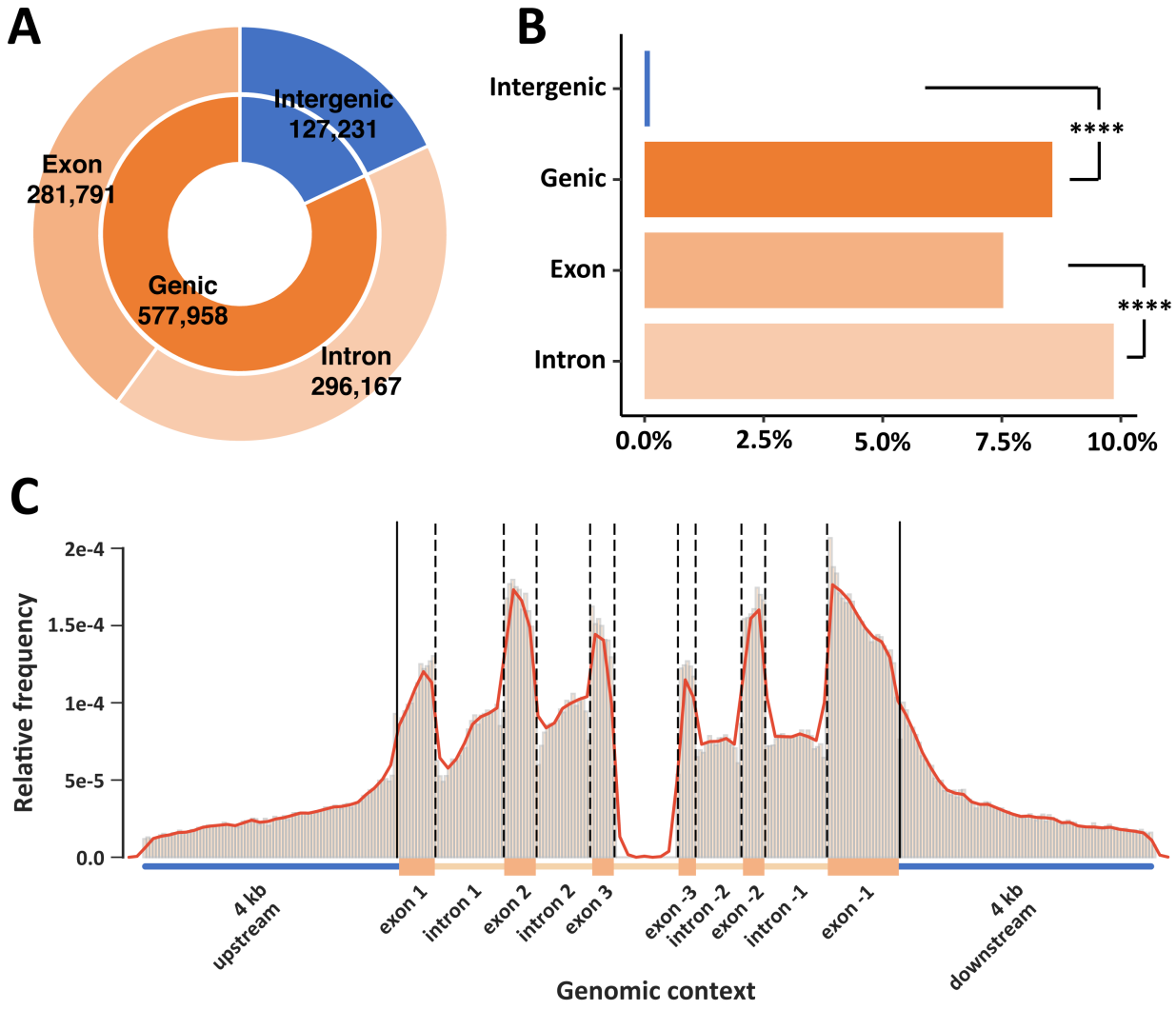


Fig. 1. DNA methylation landscape

(**A**) Distribution of methylated CpG across intergenic (18%), genic (82%), intronic (42%) and exonic (40%) regions in the Aiptasia genome. (**B**) Normalized percentage of methylated CpGs in different regions. Chi-squared test shows significant differences between intergenic and genic regions, and between exons and intron (****p<0.0001). (**C**) Relative frequencies of methylated positions across a normalized gene model.

Methylated genes are marked by H3K36me3

Analysis of methylation patterns (see above) within gene bodies showed rapidly increasing methylation levels after the transcription start site (TSS) that are maintained before slowly decreasing towards the transcription termination site (TTS) (Fig. S3A). Interestingly, we found that gene body methylation in Aiptasia is positively correlated with expression (Fig. 2A), suggesting that DNA methylation either increases the expression of genes or that DNA methylation is increased as a consequence of transcription whereby increased expression results in methylation of the respective loci. The latter interpretation would be in line with recent findings in mouse embryonic stem cells (Neri et al. 2017), which demonstrated that gene body methylation is established and maintained as a result of active transcription by RNA polymerase II (Pol II) and recruitment of the histone modifying protein SetD2 that trimethylates histone H3 at lysine 36 (H3K36me3). This histone mark is specifically bound via the PWWP domain present in the DNA methyltransferase Dnmt3b, which in turn methylates the surrounding DNA accordingly, resulting in the inhibition of transcription initiation from cryptic promoters within the gene body and thus a significant reduction of spurious transcription.

Analysis of the Aiptasia gene set identified a DNMT3 gene (AIPGENE24404) that also encodes a PWWP domain as reported for the mouse homolog. In order to test if the mechanism previously described in mice is conserved in Aiptasia, we performed a ChIP-Seq experiment using a validated antibody against H3K36me3 (Fig. S4-S5). As predicted, our analysis confirmed a significantly higher association of H3K36me3 with methylated genes ($p = 2.48 \times 10^{-20}$ for highly methylated genes and all methylated genes, Fig. 2B and C) suggesting that DNA methylation in Aiptasia might indeed be a consequence of expression. We then analyzed if methylated genes also exhibited significantly lower levels of spurious transcription in Aiptasia.

Analysis of transcriptional profiles of methylated and unmethylated genes indeed showed significantly lower levels of spurious transcription along the gene body of methylated genes ($p < 2 \times 10^{-6}$, Fig. 3A).

A dampening effect of DNA methylation on transcription was also observed with regard to transcriptional noise similar to findings in the coral *Stylophora pistillata* (Liew et al. 2017). Regression analysis of median methylation levels and the coefficient of transcriptional variation of genes showed that, given the same expression level, methylated genes always exhibited lower levels of transcriptional variation (Fig. 3B).

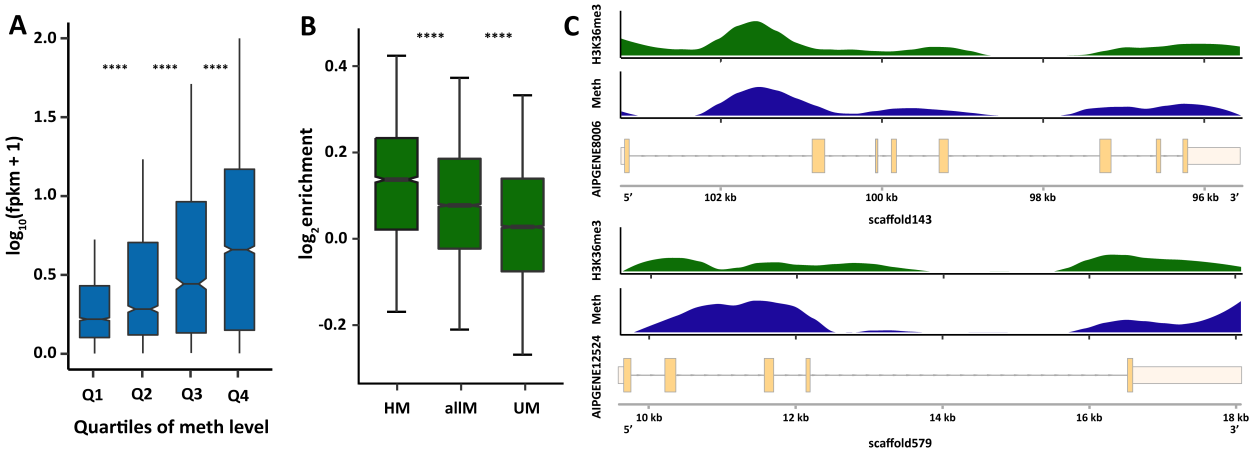


Fig. 2. DNA methylation is associated with higher expression

(**A**) Gene expression is positively correlated with median methylation levels, *t*-test *p* values are 7.65×10^{-21}, 3.75×10^{-14} and 1.75×10^{-13} for the first quartile (Q1) and the second quartile (Q2) of methylation levels, Q2 and Q3, and Q3 and Q4, respectively. (**B**) ChIP-Seq analysis of H3K36me3 signals show significant enrichment in methylated genes (*t*-test *p* values: 2.48×10^{-20} for highly methylated genes (HM) and all methylated genes (allM), and 1.06×10^{-72} for unmethylated genes (UM) and allM). Highly methylated genes show the strongest enrichment

with H3K36me3 followed by all methylated genes. In contrast unmethylated genes show only weak enrichment of H3K36me3 over input controls. (**C**) Distribution of H3K36me3 enrichment and DNA methylation levels across two exemplary gene models. H3K36me3 and DNA methylation show coinciding distribution patterns over genes.

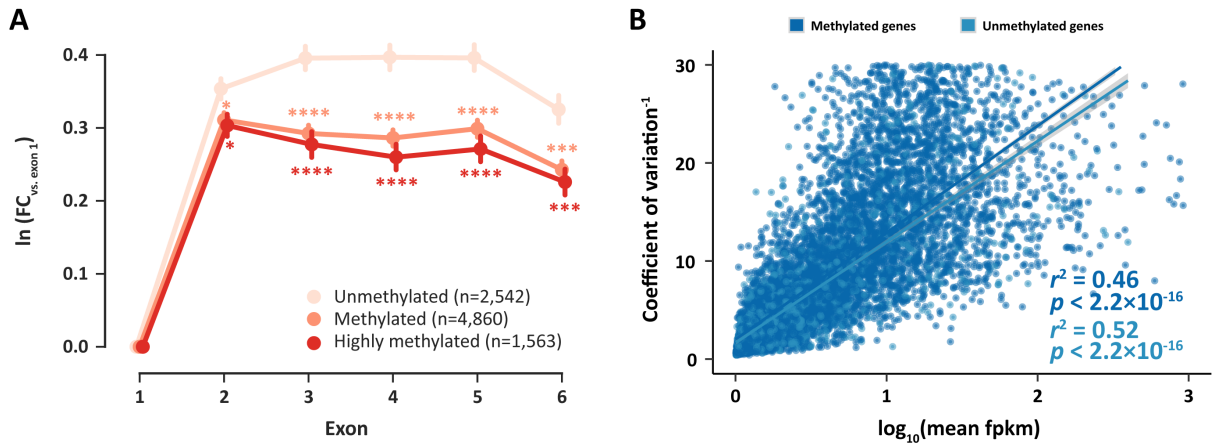


Fig. 3. DNA methylation regulates transcriptional homeostasis

(A) Spurious transcription in gene bodies is significantly lower in methylated and highly methylated genes. The *y*-axis shows the natural logarithm of the coverage fold change of exons 1–6 vs. exon 1. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; ****: $p < 0.0001$. (B) There is a linear relationship between the inverse of transcriptional noise (CV^{-1}) and log expression level (log_{10}fpkm). Given same expression level, methylated genes always show lower levels of transcriptional noise. For methylated genes, n = 8,561, $r^2 = 0.46$, $p < 2.2 \times 10^{-16}$, for unmethylated genes, n = 2,491, $r^2 = 0.52$, $p < 2.2 \times 10^{-16}$.

DNA methylation regulates transcriptional homeostasis during symbiosis

Based on our previous findings, we investigated if DNA methylation might also be involved in the regulation of symbiosis by identifying differentially methylated genes (DMGs) between symbiotic and aposymbiotic Aiptasia. Comparison of DNA methylation patterns using Principal Component Analysis (PCA) clearly separated symbiotic and aposymbiotic individuals by the first principal component, which accounted for ~18% of the variance (Fig. 4 and Fig. S6). This analysis echoed the findings from a PCA analysis on gene expression where symbiosis state was separated by the second principal component accounting for ~25% of the variance (Fig. 4B) (Venables and Ripley 2002) and highlighted that specific changes in DNA methylation patterns occurred in response to symbiosis. Analysis of DNA methylation and expression profiles using correlation analyses further confirmed this finding, providing additional evidence that the observed changes were indeed treatment specific (Fig. S6).

Subsequently we analyzed changes in DNA methylation and gene expression between symbiotic and aposymbiotic Aiptasia to assess their correlation on potential biological functions in symbiosis. We determined differentially methylated genes using a generalized linear model from Foret *et al*. (Foret et al. 2012) that was modified to allow for replicate-aware analysis. This approach identified 2,133 DMGs (FDR ≤ 0.05, Supplement Table S1) that specifically changed their methylation status in response to symbiosis. To verify these results, we sequenced a subset of 14 DMGs using bisulfite PCRs. The results show a strong correlation ($r^2 = 0.815$ and $p = 1\times10^{-5}$ for Apo, $r^2 = 0.922$ and $p = 5.2 \times 10^{-8}$ for Sym) to our WGBS and confirm the observed methylation changes within these loci (Fig. S7).

Analysis of gene expression changes in the same 12 samples (i.e., 6 symbiotic and 6

178 aposymbiotic anemones) identified 1,278 differentially expressed genes (DEGs, FDR ≤ 0.05,
179 Supplement Table S2), of which 14 genes were subsequently confirmed via qPCR (Fig. S8).
180 However, analysis of the overlap between DMGs and DEGs showed only 103 genes that were
181 shared, suggesting that differentially expressed genes are not necessarily the same cohort of
182 genes that are differentially methylated. Functional enrichment analyses based on Gene Ontology
183 (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of all DMGs and
184 DEGs identified several symbiosis relevant functions and pathways in both groups (Supplement
185 Table S3-S10).

186 Based on the finding that gene body DNA methylation is likely a consequence of active
187 transcription, we hypothesized that changes in DNA methylation patterns might also provide a
188 record of transcriptional activity over longer periods of time. We therefore tested if differential
189 methylation and acute transcriptional changes, obtained from our RNA-Seq analysis, provide a
190 complementary view of the processes underlying symbiosis. For this we compared enrichment of
191 symbiosis-specific pathways across the sets of 2,133 DMGs, 1,278 DEGs, and the combined set
192 of both DMGs and DEGs (3,308 DEMGs). Interestingly, we observed that the combined data set
193 (DEMGs) provided significantly lower p-values for previously identified symbiosis-related
194 pathways, including apoptosis, phagosome formation, nitrogen metabolism, and arginine
195 biosynthesis, among others (paired *t*-test: DEMG vs. DEG $p = 0.015$; DEMG vs. DMG $p = 0.009$)
196 (Fig. 4C and Supplement Table S11). This suggested that changes in methylation and
197 transcription indeed provide complementary information with regard to transcriptional
198 adjustments in response to symbiosis.

199

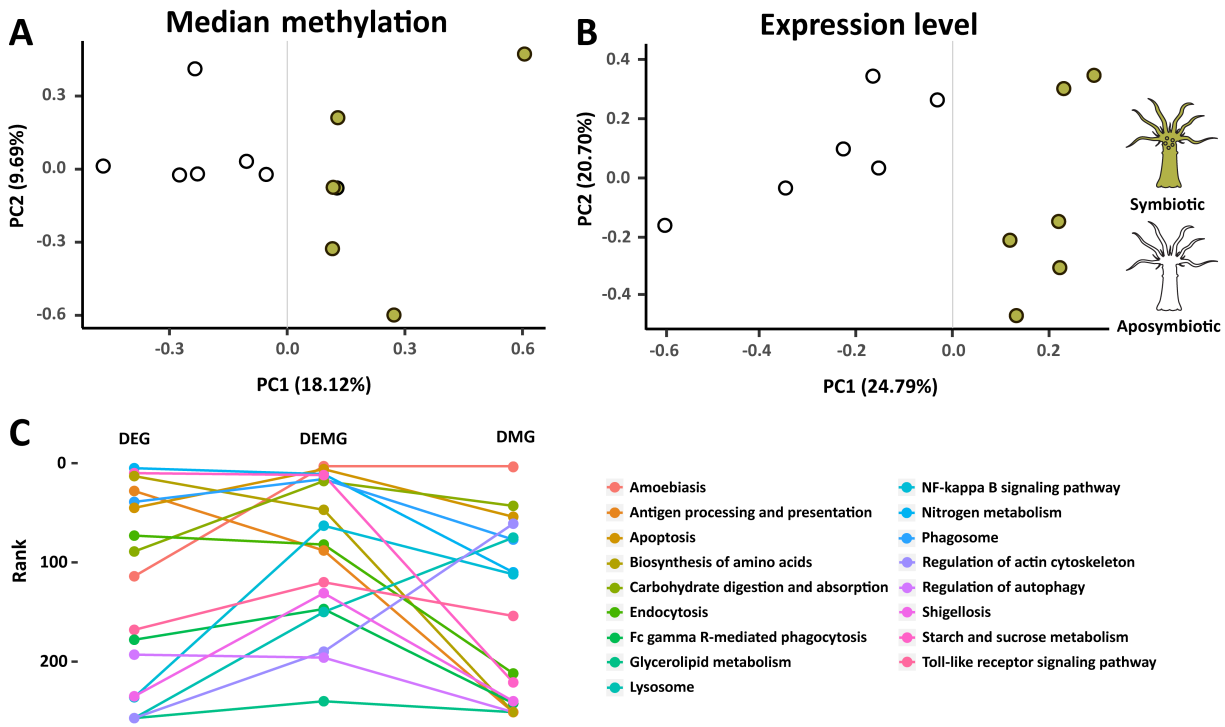


Fig.4. PCA and KEGG pathway enrichment analysis

(**A**, **B**) PCA (Principal Component Analysis) of gene expression and median methylation levels of Aiptasia genes. Both gene expression and DNA methylation separate samples by symbiosis state. (**C**) KEGG pathway enrichment analysis. The combined sets of differentially expressed and differentially methylated genes (DEMG) provides significant lower *p* values (front ranks) for symbiosis related pathways.

DMGs and DEGs are involved in all stages of symbiosis

Analysis of the combined DMG and DEG gene set showed significant enrichment of genes involved in the distinct phases of symbiosis, that is symbiosis establishment, maintenance, and breakdown (Davy et al. 2012). Using an integrated pathway analysis based on known

212 molecular interactions between proteins we found that these processes are linked through several
213 DMGs and/or DEGs (Fig. S9 and Fig. S10, and see Supplement Table S11-S12 and
214 Supplementary discussion).

215 For instance, we found numerous symbiosis-related receptors to respond to symbiosis on
216 a transcriptional and/or methylation level (Fig. S9), including C-type lectins (Fig. S9.3), Toll-like
217 receptors (Fig. S9.5), and the scavenger receptor SRB1 (Fig. S9.2) that has previously been
218 implicated in symbiont recognition in the sea anemone *Anthopleura elegantissima* (Rodriguez-
219 Lanetty et al. 2006; Neubauer et al. 2016). Following symbiont recognition, we also found
220 several known engulfment and sorting-related genes to change in methylation and/or expression
221 such as Rab5 (Fig. S9.10), sorting nexin (Fig. S9.17), Rac1 (Fig. S9.6), the lysosomal-associated
222 membrane protein 1/2 (Fig. S9.22), and many genes related to the cytoskeleton and movement
223 (Fig. S9.33-39).

224 As expected in a metabolic symbiosis (Muscatine 1990; Davy et al. 2012) we also identified a
225 large number of genes involved in nutrient exchange. These included genes involved in the
226 provision of inorganic carbon in the form of CO_2 or bicarbonate (HCO_3^-) to fuel symbiont driven
227 photosynthesis (Rädecker et al. 2017) (Fig. S10.1) as well as genes involved in the exchange of
228 fixed carbon in the form of lipids (Fig. S10.11), sugars and amino acids (Fig. S10.10, S10.4)
229 (Oakley et al. 2016). Concordantly, we also found that genes involved in nitrogen acquisition,
230 such as ammonium transporter (Fig. S10.2) and genes involved in glutamate metabolism (Fig.
231 S10.5-7), respond to symbiosis.

232 Finally, our analysis also highlighted genes putatively involved in the expulsion or
233 degradation of symbionts in response to environmental stress or as a means to control symbiont

densities. Autophagy is of interest in this regard because it links to other membrane trafficking pathways and to apoptosis, and evidence suggests that autophagy also plays a role in removal of symbionts during bleaching (Dunn et al. 2007; Downs et al. 2009). Intracellular degradation of the symbiont is a result of reengagement of the phagosomal maturation process or autophagic digestion of the symbiont by the host cell (Davy et al. 2012), and we find both apoptosis- and autophagy-related genes to significantly change in their methylation and/or expression level. These include the apoptosis genes RAC serine/threonine-protein kinase (Fig. S9.25), Caspase 7 (Fig. S9.31), Caspase 8 (CASP8) (Fig. S9.30), Nitric oxide synthase (Fig. S9.21) and Bcl2 (Fig. S9.27), as well as the Autophagy proteins 5 and 10 (Fig. S9.14-15), among others.

Discussion

To assess the role of CpG methylation in the cnidarian-dinoflagellate symbiosis, we undertook a global analysis of changes in the DNA methylomes and transcriptomes of aposymbiotic and symbiotic Aiptasia. In contrast to their vertebrate counterparts, only 6.37% of the CpGs in the Aiptasia genome are methylated, but their distribution is highly non-random ($p < 3 \times 10^{-300}$) and that methylated CpGs are most highly localized in gene bodies (18.4% of CpGs). Analysis of the distribution of the histone modification H3K36me3 further showed significant enrichment of this epigenetic mark in methylated genes, echoing findings in mammals and invertebrates (Nanty et al. 2011). More importantly, we find that methylated genes show significant reduction of spurious transcription and transcriptional noise (Fig. 2B), suggesting that both the underlying mechanism of epigenetic crosstalk as well as the biological function of DNA methylation is evolutionarily conserved throughout metazoans. These results highlight a tight interaction of transcription and epigenetic mechanisms in optimizing gene expression in response

257 to changing transcriptional needs (Neri et al. 2017). Further support for such a role is provided
258 by the analysis of differentially methylated and differentially expressed genes, which, when
259 combined, showed significant increase in enrichment of symbiosis relevant processes. This
260 suggests that DNA methylation and transcriptome analyses provide complementary views of
261 cellular responses to symbiosis whereby methylation changes provide a transcriptional record of
262 longer-term transcriptional adjustments.

263 While our analysis identified several genes, processes, and pathways previously reported
264 to be involved in symbiosis, it further highlights their intricate molecular interactions. Symbiosis
265 recognition, sorting and breakdown are interconnected processes, which is reflected in the
266 observed changes in methylation and expression. The molecular machinery involved in
267 phagosome maturation is tightly linked to autophagy and apoptosis enabling the host to respond
268 to potential pathogen invasion but also to degrade and remove dead or unsuitable symbionts.
269 This is strongly supported by immunofluorescence examinations of *Aiptasia pulchella*
270 gastrodermal cell macerates, showing that Rab5 appears around healthy, newly ingested and
271 already established *Symbiodinium*, but is replaced by Rab7 in heat-killed or DCMU-treated
272 newly ingested *Symbiodinium*. Conversely, Rab7 is absent from untreated newly infected or
273 already-established *Symbiodinium* (Chen et al. 2003; Chen et al. 2004).

274 Rab5 is also required for the exosomal release of CD63 (Baietti et al. 2012), which
275 mediates the endocytotic sorting process and transport to lysosomes (Latysheva et al. 2006). This
276 process is further regulated by Rac1 (Anitei et al. 2010) in conjunction with sorting nexin and the
277 GTPase Rho, all of which were also identified in our analyses. The sorting of phagocytosed
278 *Symbiodinium* is critical to symbiosis establishment as *Symbiodinium* cells are phagocytosed at
279 the apical end and transported to the base of the cell, where they are protected from digestion. In

280 contrast, *Symbiodinium* staying at the apical end of the cell are degraded (McAuley and Smith
281 1982).

282 Similar to the processes of symbiosis initiation and breakdown, we also found significant
283 enrichment of genes involved in nutrient exchange and many of these transporters have
284 previously been implicated in symbiosis maintenance (Davy et al. 2012; Lin et al. 2015).
285 Notably, this also included genes involved in the transport and assimilation of ammonium.
286 Nitrogen is a main limiting nutrient in coral reefs (Cook et al. 1992; Grover et al. 2008; Radecker
287 et al. 2015), and the coral-dinoflagellates symbiosis has been proposed to increase the efficiency
288 of nitrogen utilization by both partners (Wang and Douglas 1998b) whereby the underlying
289 nature of this mechanism is currently debated (Wang and Douglas 1998a; Aranda et al. 2016).

290

291 **Conclusions**

292 This study provides the first analysis of the function and role of DNA methylation in a symbiotic
293 anthozoan. Our results show that the epigenetic crosstalk between the histone mark H3K36me3
294 and gene body methylation is conserved in cnidarians and reveal a role of gene body methylation
295 in reducing of spurious transcription and transcriptional noise. Furthermore, we show that
296 changes in DNA methylation patterns are specific to symbiosis and imply a functional in the
297 establishment, maintenance, and breakdown of this important symbiotic association. Our
298 findings therefore provide evidence for a role of DNA methylation as an epigenetic mechanism
299 involved in the maintenance of transcriptional homeostasis during the cnidarian-dinoflagellate
300 symbiosis. The premise that epigenetic mechanisms play a role in organismal acclimation
301 warrants future experiments targeted to investigate if DNA methylation could also contribute to

resilience through the epigenetic adjustment of transcription in response to environmental stress in Aiptasia and corals.

Data availability

Sequencing data of Bis-Seq, RNA-Seq and ChIP-Seq were deposited in NCBI Sequence Read Archive (SRA) under BioProject codes PRJNA415358.

Acknowledgements

Research reported in this publication was supported by funding from King Abdullah University of Science and Technology (KAUST). We thank Prof. John Pringle for the provision of the initial Aiptasia CC7 and *Symbiodinium* SSB01 cultures. We thank Sebastian Baumgarten for his help with establishing the Aiptasia cultures and critical reading of the manuscript.

Author contributions

M.A. conceived and coordinated the project. Y.L., G.C., M.J.C. and N.Z. performed experiments. M.A., C.R.V. and Y.J.L. provided tools and/or data. C.T.M. constructed libraries for whole genome bisulfite sequencing, ChIP-Seq and RNA-Seq. Y.L., Y.J.L. and M.A. analyzed expression, methylation and ChIP-Seq data. M.A. and Y.L. wrote the manuscript with input from Y.J.L. and C.R.V. All authors read and approved the final manuscript.

References and Notes:

1. Spalding MD, Grenfell AM. New estimates of global and regional coral reef areas. *Coral Reefs* **16**, 225-230 (1997).
2. Davidson MG. Protecting coral reefs: The principal national and international legal instruments. *The Harvard Environmental Law Review* **26**, 48 (2002).
3. Sylvain JC. How to Protect A Coral Reef: The Public Trust Doctrine and the Law of the Sea. *Sustainable Development Law & Policy* **7**, 4 (2006).
4. Davy SK, Allemand D, Weis VM. Cell biology of cnidarian-dinoflagellate symbiosis. *Microbiology and molecular biology reviews : MMBR* **76**, 229-261 (2012).
5. Meyer E, Weis VM. Study of cnidarian-algal symbiosis in the "omics" age. *The Biological bulletin* **223**, 44-65 (2012).
6. Lämke J, Bäurle I. Epigenetic and chromatin-based mechanisms in environmental stress adaptation and stress memory in plants. *Genome Biology* **18**, 124 (2017).
7. Rando OJ, Verstrepen KJ. Timescales of Genetic and Epigenetic Inheritance. *Cell* **128**, 655-668 (2007).
8. He X-J, Chen T, Zhu J-K. Regulation and function of DNA methylation in plants and animals. *Cell Res* **21**, 442-465 (2011).
9. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* **13**, 484-492 (2012).
10. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet* **9**, 465-476 (2008).
11. Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev* **16**, 6-21 (2002).
12. Richards EJ. Inherited epigenetic variation - revisiting soft inheritance. *Nat Rev Genet* **7**, 395-401 (2006).
13. Neri F*, et al.* Intragenic DNA methylation prevents spurious transcription initiation. *Nature* **543**, 72-77 (2017).
14. Torda G*, et al.* Rapid adaptive responses to climate change in corals. *Nature Clim Change* **7**, 627-636 (2017).
15. Baumgarten S*, et al.* The genome of Aiptasia, a sea anemone model for coral symbiosis. *Proc Natl Acad Sci U S A* **112**, 11893-11898 (2015).
16. Voolstra CR. A journey into the wild of the cnidarian model system Aiptasia and its symbionts. *Molecular Ecology* **22**, 4366-4368 (2013).

17. Tucker KL. Methylated cytosine and the brain: a new base for neuroscience. *Neuron* **30**, 649-652 (2001).
18. Liew YJ*, et al.* Epigenome-associated phenotypic acclimatization to ocean acidification in a reef-building coral. *bioRxiv*, (2017).
19. Dixon GB, Bay LK, Matz MV. Patterns of gene body methylation predict coral fitness in new environments. *bioRxiv*, (2017).
20. Wang X*, et al.* Function and Evolution of DNA Methylation in *Nasonia vitripennis*. *PLoS Genet* **9**, e1003872 (2013).
21. Feng S*, et al.* Conservation and divergence of methylation patterning in plants and animals. *Proc Natl Acad Sci U S A* **107**, 8689-8694 (2010).
22. Krzywinski M*, et al.* Circos: an information aesthetic for comparative genomics. *Genome Res* **19**, 1639-1645 (2009).
23. Venables WN, Ripley BD. Exploratory Multivariate Analysis. In: *Modern Applied Statistics with S* (ed^(eds). 4 edn. Springer-Verlag New York (2002).
24. Foret S*, et al.* DNA methylation dynamics, metabolic fluxes, gene splicing, and alternative phenotypes in honey bees. *Proc Natl Acad Sci U S A* **109**, 4968-4973 (2012).
25. Rodriguez-Lanetty M, Phillips WS, Weis VM. Transcriptome analysis of a cnidarian-dinoflagellate mutualism reveals complex modulation of host gene expression. *BMC Genomics* **7**, 23 (2006).
26. Muscatine L. *The role of symbiotic algae in carbon and energy flux in reef corals*. Elsevier Science Publishing Company (1990).
27. Rädecker N, Pogoreutz C, Wild C, Voolstra CR. Stimulated Respiration and Net Photosynthesis in *Cassiopeia* sp. during Glucose Enrichment Suggests *in hospite* CO_2 Limitation of Algal Endosymbionts. *Frontiers in Marine Science* **4**, (2017).
28. Oakley CA, Ameismeier MF, Peng L, Weis VM, Grossman AR, Davy SK. Symbiosis induces widespread changes in the proteome of the model cnidarian Aiptasia. *Cellular Microbiology* **18**, 1009-1023 (2016).
29. Dunn SR, Schnitzler CE, Weis VM. Apoptosis and autophagy as mechanisms of dinoflagellate symbiont release during cnidarian bleaching: every which way you lose. *Proceedings of the Royal Society B: Biological Sciences* **274**, 3079-3085 (2007).
30. Downs CA*, et al.* Symbiophagy as a cellular mechanism for coral bleaching. *Autophagy* **5 2**, 211-216 (2009).

31. Nanty L*, et al.* Comparative methylomics reveals gene-body H3K36me3 in Drosophila predicts DNA methylation and CpG landscapes in other invertebrates. *Genome Research* **21**, 1841-1850 (2011).
32. Chen MC, Cheng YM, Sung PJ, Kuo CE, Fang LS. Molecular identification of Rab7 (ApRab7) in *Aiptasia pulchella* and its exclusion from phagosomes harboring zooxanthellae. *Biochem Biophys Res Commun* **308**, 586-595 (2003).
33. Chen MC, Cheng YM, Hong MC, Fang LS. Molecular cloning of Rab5 (ApRab5) in *Aiptasia pulchella* and its retention in phagosomes harboring live zooxanthellae. *Biochem Biophys Res Commun* **324**, 1024-1033 (2004).
34. Baietti MF*, et al.* Syndecan-syntenin-ALIX regulates the biogenesis of exosomes. *Nat Cell Biol* **14**, 677-685 (2012).
35. Latysheva N*, et al.* Syntenin-1 is a new component of tetraspanin-enriched microdomains: mechanisms and consequences of the interaction of syntenin-1 with CD63. *Mol Cell Biol* **26**, 7707-7718 (2006).
36. Anitei M*, et al.* Protein complexes containing CYFIP/Sra/PIR121 coordinate Arf1 and Rac1 signalling during clathrin-AP-1-coated carrier biogenesis at the TGN. *Nat Cell Biol* **12**, 330-340 (2010).
37. McAuley PJ, Smith DC. The Green Hydra Symbiosis. V. Stages in the Intracellular Recognition of Algal Symbionts by Digestive Cells. *Proceedings of the Royal Society B: Biological Sciences* **216**, 7-23 (1982).
38. Lin S*, et al.* The Symbiodinium kawagutii genome illuminates dinoflagellate gene expression and coral symbiosis. *Science* **350**, 691-694 (2015).
39. Cook CB, Mullerparker G, Delia CF. Ammonium Enhancement of Dark Carbon Fixation and Nitrogen Limitation in Symbiotic Zooxanthellae - Effects of Feeding and Starvation of the Sea Anemone *Aiptasia Pallida*. *Limnology and Oceanography* **37**, 131-139 (1992).
40. Grover R, Maguer J-F, Allemand D, Ferrier-Pagès C. Uptake of dissolved free amino acids by the scleractinian coral *Stylophora pistillata*. *Journal of Experimental Biology* **211**, 860-865 (2008).
41. Radecker N, Pogoreutz C, Voolstra CR, Wiedenmann J, Wild C. Nitrogen cycling in corals: the key to understanding holobiont functioning? *Trends Microbiol* **23**, 490-497 (2015).
42. Wang J, Douglas AE. Nitrogen recycling or nitrogen conservation in an alga-invertebrate symbiosis? *J Exp Biol* **201 (Pt 16)**, 2445-2453 (1998).
43. Wang J, Douglas A. Nitrogen recycling or nitrogen conservation in an alga-invertebrate symbiosis? *The Journal of Experimental Biology* **201**, 2445-2453 (1998).

44. Aranda M*, et al.* Genomes of coral dinoflagellate symbionts highlight evolutionary adaptations conducive to a symbiotic lifestyle. *Scientific Reports* **6**, 39734 (2016).

411

412 Supplementary Materials for

413

414 # DNA methylation regulates transcriptional homeostasis of algal endosymbiosis

415 # in the coral model Aiptasia

416 Yong Li[1], Yi Jin Liew[1], Guoxin Cui[1], Maha J. Cziesielski[1], Noura Zahran[1], Sebastian

417 Baumgarten[1], Craig T. Michell[1], Christian R. Voolstra[1], Manuel Aranda[1]

418 [1]Red Sea Research Center, King Abdullah University of Science and Technology (KAUST),

419 Thuwal, KSA

420

421 correspondence to: manuel.aranda@kaust.edu.sa

422

423 ## Materials and methods

424 ### *Exaiptasia pallida* Culture and DNA/RNA Extraction

425 *Exaiptasia pallida* of the clonal strain CC7 (Sunagawa et al. 2009) was used for this study.
426 Anemones were maintained in polycarbonate tubs with autoclaved seawater at ~25 °C on a 12 h:
427 12 h light: dark cycle at 20-40 µmol m^{-2} s^{-1} light intensity and fed freshly hatched *Artemia*
428 *nauplii* (brine-shrimp) approximately twice per week. To generate aposymbiotic anemones,
429 animals were subjected to multiple cycles of cold-shock treatment and the photosynthesis
430 inhibitor diuron (Sigma-Aldrich, St. Louis, MO) as described in Baumgarten (2015)
431 (Baumgarten et al. 2015). Aposymbiotic anemones were kept individually in 15 ml autoclaved

seawater in 6-well plates and inspected by fluorescence stereomicroscopy to confirm complete absence of dinoflagellates. In order to exclude potential batch effects as source of DNA methylation changes we first generated four separate batches of aposymbiotic anemones and maintained them for a period of 1 year before beginning of the experiment described below.

To generate symbiotic anemones, we then separately infected aposymbiotic CC7 individuals from each of the four aposymbiotic cultures described above using the compatible Clade B *Symbiodinium* strain SSB01, originally isolated from Aiptasia strain H2 (Xiang et al. 2013; Baumgarten et al. 2015). The four batches of symbiotic anemones were maintained for further 12 months under regular culture condition as described above. The corresponding four aposymbiotic cultures were maintained in darkness until 3 months before collection. For the last 3 months individuals from these aposymbiotic cultures were subjected to the same culture conditions as the symbiotic cultures in order to monitor for unwanted spontaneous re-establishment of symbiosis under light.

After the 12-month experimental period, we collected six biological replicates from each of the four aposymbiotic and symbiotic cultures (one additional replicate was taken from batches 1 and 2 of each treatment) for subsequent DNA and RNA extraction as described below.

For each treatment, 6 biological replicates, weighing 20-28 mg (wet weight), were extracted using the AllPrep DNA/RNA/miRNA Universal Kit (Qiagen, Hilden, Germany). The manufacturer's protocol was followed with the omission of the optional step 4 (temporal storage at 4°C if not performing DNA purification immediately). DNA concentrations were determined using a Qbit dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA). RNA concentrations and integrity were determined using a Bioanalyzer Nano RNA Kit (Agilent Technologies, Santa Clara, CA).

455

456 **RNA-Seq and Bisulfite Sequencing**

457 Directional mRNA libraries were produced using the NEBNext® Ultra™ Directional RNA
458 Library Prep Kit for Illumina® (NEB) following the manufacturer's protocol.

459 Bisulfite DNA libraries were prepared following a modified version of the NEBNext®
460 Ultra™ II DNA Library Prep Kit for Illumina® (NEB). Methylated TruSeq Illumina® adapters
461 (Illumina) were used during the adapter ligation step followed by bisulfite conversion with the
462 EpiTect Bisulfite kit (QIAGEN), with the following cycling conditions (95°C – 5 min, 60°C – 25
463 min, 95°C – 5 min, 60°C – 85 min, 95°C – 5 min, 60°C – 175 min, and 3 cycles of 95°C – 5 min,
464 60°C – 180 min, hold at 20°C ≤ 5 hours) (reference is Illumina Bisulfite).

465 The final libraries were enriched with the KAPA HiFi HotStart Uracil+ ReadyMix (2X)
466 (KAPA Biosystems) following the standard protocol for bisulfite-converted NGS library
467 amplification. Final libraries were quality checked using the Bioanalyzer DNA 1K chip
468 (Agilent), and quantified using Qubit 2.0 (Thermo Fisher Scientific), and then pooled in
469 equimolar ratios and sequenced on the HiSeq2000.

470

471 **Identification of methylated CpGs**

472 Sequencing of the 12 libraries (2 conditions, 6 biological replicates each) resulted in 819
473 million read pairs from 8 lanes of the Illumina HiSeq2000 platform. Adapters were trimmed
474 from the raw sequences using cutadapt v1.8 (Martin 2011). Subsequently, trimmed reads were
475 mapped to the *Exaiptasia pallida* genome (Baumgarten et al. 2015) using Bowtie2 v2.2.3
476 (Langmead and Salzberg 2012), and methylation calls was performed using Bismark v0.13

(Krueger and Andrews 2011).

Three filters were used to reduce false positives. Firstly, for each position with k methylated reads mapping to it, the probability of it occurring through sequencing error (i.e. unmethylated position appearing as methylated) was modelled using a binomial distribution B(n, p), where n is the coverage (methylated + unmethylated reads) and p the probability of sequencing error (set to 0.01). We kept positions with k methylated reads if $P(X \geq k) < 0.05$ (post-FDR correction). Secondly, retained methylated positions had to have ≥ 1 methylated read in all six biological replicates of at least one growth condition. Finally, median coverage of retained positions across all 12 samples had to be ≥ 10.

Assignment of genomic context to methylated cytosines

Based on the gene annotation of the *Exaiptasia pallida* genome (GFF3 file) (Baumgarten et al. 2015) and the positional coordinates of the methylated cytosines produced by Bismark, we annotated every methylated cytosine based on the genomic context, including whether the methylated position resides in a genic or intergenic region, and the distances to the 5' and 3' end of each genomic feature (gene/intergenic region/exon/intron).

CpG bias

Methylated cytosines are frequently spontaneously deaminated to uracil which can be subsequently converted to thymine after DNA repair. As a result of this process, methylated CpGs are expected to decrease in abundance over evolutionary time, and the ratio of observed to expected CpGs (CpG O/E) has previously been used to predict putatively methylated and

499 unmethylated genes (Suzuki et al. 2007; Wang and Leung 2008). CpG O/E of *Exaiptasia pallida*
500 protein coding genes were calculated according to J. Zeng *et al* (Zeng and Yi 2010).

501

502 **Identification of differentially methylated genes**

503 Using the methylation level of aposymbiotic genes as a control, generalized linear models
504 (GLMs) (Hastie and Pregibon 1992) were implemented in R (R Core Team 2016) to identify
505 genes that were differentially methylated in the symbiotic treatment. The general formula used
506 was:

507 glm(methylated, non_methylated ~ treatment * position, family="binomial")

508 where "methylated, non_methylated" was a two-column response variable denoting the
509 number of methylated and non-methylated reads at a particular position. For predictor variables,
510 "position" denoted relative position of the methylated site in the gene, while "treatment" denoted
511 symbiotic or aposymbiotic conditions. Data from individual replicates were entered separately to
512 assign equal weightage to each replicate, as pooling results in a disproportionate skew towards
513 the replicate with the highest coverage. Genes with < 5 methylated positions were filtered out to
514 reduce type I errors; and genes with FDR ≤ 0.05 were considered as differentially methylated
515 genes (DMGs).

516

517 **Identification of differentially expressed genes**

518 RNA-Seq generated 889 million raw read pairs from six lanes on the Illumina Hiseq2000
519 platform. Adaptors, primers and low quality bases were removed from the ends of raw reads
520 using Trimmomatic v0.33 (ILLUMINACLIP:TruSeq2-PE.fa:4:25:9 LEADING:28

TRAILING:28 SLIDINGWINDOW:4:30 MINLEN:50). The resulting trimmed reads were mapped to the *Exaiptasia pallida* genome using HISAT v2.0.1 (Kim et al. 2015) and transcripts were assembled based on the *Exaiptasia pallida* gene models (GFF3 file) using StringTie v1.2.2 (Pertea et al. 2015). Trinity (align_and_estimate_abundance.pl – Bowtie2 v2.2.7/RSEM v1.2.22/edgeR v3.10.5) (Robinson et al. 2010; Grabherr et al. 2011; Li and Dewey 2011; Langmead and Salzberg 2012; Haas et al. 2013) was run against the transcripts using trimmed reads for expression abundance estimation, then differentially expressed genes (DEGs) were identified with FDR ≤ 0.05.

PCA and correlation matrices

Median methylation levels and log FPKM (base 2) of genes were shifted to be zero centered and analyzed by Principal Component Analysis (PCA) using the prcomp function in R.

Using the same data we calculated correlation matrices (Pearson correlation coefficient) and clustered samples with hclust implemented in R using complete linkage and euclidean distance .

Spurious transcription analysis

Trimmed reads were mapped to the *Exaiptasia pallida* genome using HISAT2 v2.1.0 and mapping coverage per position was extracted using BEDtools v2.17.0. Coverage per exon was calculated and normalized across all 6 replicates (assuming every replicate had 1 million coverage in total), then average coverage ratios of exon 2 to 6 versus exon 1 per gene were calculated to determine spurious transcription levels.

544

545 **GO enrichment of DMGs and DEGs**

546 GO (Gene Ontology) (Ashburner et al. 2000) annotation was based on the previously
547 published *Exaiptasia pallida* genome (Baumgarten et al. 2015). Functional enrichment of DMGs
548 and DEGs was carried out with topGO respectively (Adrian Alexa 2016) using default settings.
549 GO terms with $p \leq 0.05$ were considered significant, and the occurrence of at least ≥ 5 times in
550 the background set was additionally required for DMGs. Multiple testing correction was not
551 applied to the resulting *p*-values as the tests are considered to be non-independent (Adrian Alexa
552 2016).

553

554 **KEGG enrichment of DMGs and DEGs**

555 KEGG (Kyoto Encyclopedia of Genes and Genomes)(Kanehisa and Goto 2000; Kanehisa et
556 al. 2016) orthology (KO) annotation was carried out by combining the KEGG annotations
557 provided in the original *Exaiptasia pallida* genome publications and a separate set of annotations
558 based on the KAAS (KEGG Automatic Annotation Server, http://www.genome.jp/tools/kaas/)
559 (parameters: GHOSTZ, Eukaryotes, Bi-directional Best Hit) (Moriya et al. 2007). A KEGG
560 pathway enrichment analysis of both DMGs and DEGs was carried out using Fisher's exact test
561 and pathways with $p \leq 0.05$ were considered significant.

562

563 **Validation of gene expression changes from RNA-Seq by qPCR**

564 Three randomly picked RNA libraries per treatment were used for qPCR validation of
565 RNA-Seq results. cDNA was synthesized using Invitrogen SuperScript III First-Strand Synthesis
566 SuperMix kit. A total of 14 genes were validated for differential expression using qPCR
567 (Supplement Table S13-S15). RPS7, RPL11 and NDH5 were used as internal reference
568 standards (Lehnert et al. 2014). qPCR was carried out using Invitrogen Platinum SYBR Green
569 qPCR SuperMix-UDG kit on Applied Biosystems 7900HT Fast Real-Time PCR System. All
570 protocols were strictly followed.

571

572 **Validation of methylation changes using bisulfite PCR**

573 Three randomly picked DNA libraries per treatment were used for methylation validation.
574 Bisulfite conversion was done using the EZ-96 DNA Methylation-Gold Kit (Zymo Research). 18
575 genes were used to design primers, 14 of 18 obtained effective amplifications (Supplement Table
576 S16), then the fragments were enriched by PCR amplification using Promega PCR Master Mix.
577 Sequencing indices were added to enriched fragments using Illumina Nextera XT Index Kit.
578 Enriched fragments were sequenced on the Illumina MiSeq platform. All protocols were strictly
579 followed. 1,870x data per replicate were obtained, methylated CpGs were identified using
580 Bismark as described above. The correlations between whole genome bisulfite conversion and
581 bisulfite PCR were calculated using generalized linear model.

582

583 **Chromatin Immunoprecipitation – ChIP**

584 We used the Zymo-Spin ChIP Kit to conduct histone bound chromatin extraction, with
585 minor adjustments to manufacturer's protocol. Briefly, three biological replicates, each

586 consisting of two symbiotic anemones, were used for this experiment. Each anemone was first
587 washed with PBST (phosphate-buffered saline with 0.1% triton). Anemones were then fixed in
588 1X PBS with 1% formaldehyde for 15 minutes. To stop cross-linking reactions glycine was
589 added to the solution and left to rest for 10 more minutes. Following manufacturer's protocol, we
590 centrifuged and washed whole anemones. We prepared the Nuclei Prep Buffer according to
591 protocol and crushed the two anemones of each replicate together using a douncer for
592 homogenization. Samples were then sonicated for 15 cycles on ice (15 sec ON, 30 sec cooling)
593 to ensure fragmentation to 200-500 bp. Thereafter the protocol was followed without further
594 modifications.

595 Immunoprecipitation was achieved using a target specific antibody to histone 3 lysine 36
596 tri-methylation (H3K36me3) (ab9050, Abcam), which has been validated in many eukaryotic
597 species, including mouse(Soboleva et al. 2017), *Arabidopsis thaliana*(Wollmann et al. 2017),
598 yeast(Janke et al. 2017) and zebrafish(Vastenhouw et al. 2010; Wu et al. 2011) *et al*. Comparison
599 of Aiptasia histone H3 to the respective homologs from several species for which this antibody
600 has been previously validated showed high conservation of the N-terminal tail containing
601 position H3K36 (Fig. S4) whereby 100% conservation to the zebrafish homolog was observed.

602 Corresponding input controls for each of the 3 replicates were generated as suggested by
603 the manufacturer. DNA fragment quality and quantity were confirmed using High Sensitivity
604 DNA Reagents (Agilent Technologies, California, United States) on a bioanalyzer, after which
605 ChIP libraries were constructed using the NEBNext® ChIP-Seq Library Prep Master Mix Set
606 (#E7645, New England Biolabs, Massachusetts, United States).

607 Sequencing resulted in 10M read pairs per replicate. These read pairs were trimmed using
608 Trimmomatic and mapped to the *Exaiptasia pallida* genome using HISAT2 as described above.

609 H3K36me3 enrichments were calculated as $\log_2$(average signal/average input control) for all
610 genes, unmethylated genes and highly methylated genes (methylation level > 70 and methylation
611 density > 40). P-values were calculated using t-test.
612

613 **Antibody affinity validation through Western blotting**

614 Total protein was extracted from a snap-frozen anemone crushed in 10% TCA
615 (Trichloroacetic acid). The homogenized sample was left to incubate overnight at -20 °C to allow
616 proteins to precipitate. The solution was centrifuged at 20,000g at 4 °C for 20 minutes to collect
617 suspended proteins. The pellet was then washed three times in 80% acetone and then spun down
618 again as previously. The final pellet was then air-dried for 10-15 minutes to remove residual
619 acetone. The final protein was suspended in urea lysis buffer (7 M urea, 2 M thiourea, pH 7.5) by
620 vortexing for 2 hours.

621 Samples were then prepared for western blot by adding 4x sampling buffer (0.38 M Tris
622 base, 8% SDS, 4mM EDTA, 40% glycerol, 4mg/ml bromphenol blue) to a final concentration of
623 1X. After a 2 minute incubation at 90°C samples were ready to be run on a gel at 10-12 mA. The
624 gel was transferred to a PVDF nitrocellulose membrane, rinsed with TBS buffer (150mM NaCl,
625 25mM Tris pH7.4, 0.1% Triton X-100) and blocked for 30 min at RT in TBS containing 5% fat-
626 free powder milk. The primary antibody was diluted in TBS/milk and incubated on an undulating
627 orbital shaker overnight at 4 °C. After three washes in TBS for 10 minutes each, the membrane
628 was again blocked in TBS/milk for 20 minutes at RT before proceeding with secondary antibody
629 staining. The horseradish peroxidase-linked-antibody (Anti-Rabbit IgG HRP conjugate W4011
630 and Anti-Mouse IgG HRP conjugate W4031, Promega. Wisconsin, United States) was diluted in

631 TBS/milk (1:10000) and incubated for 2 hours at RT. After final triplicate 10 minute washes in

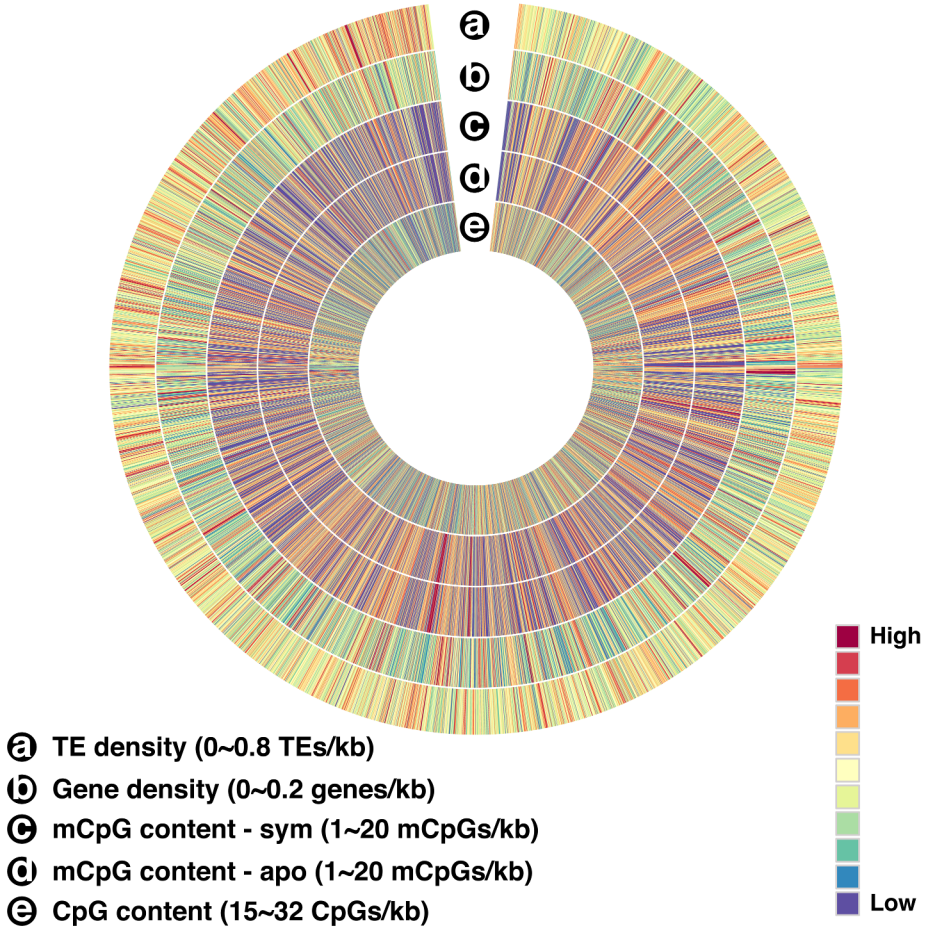
632 TBS, membranes were developed.

633

634

635 **Supplementary Figures**

636



637

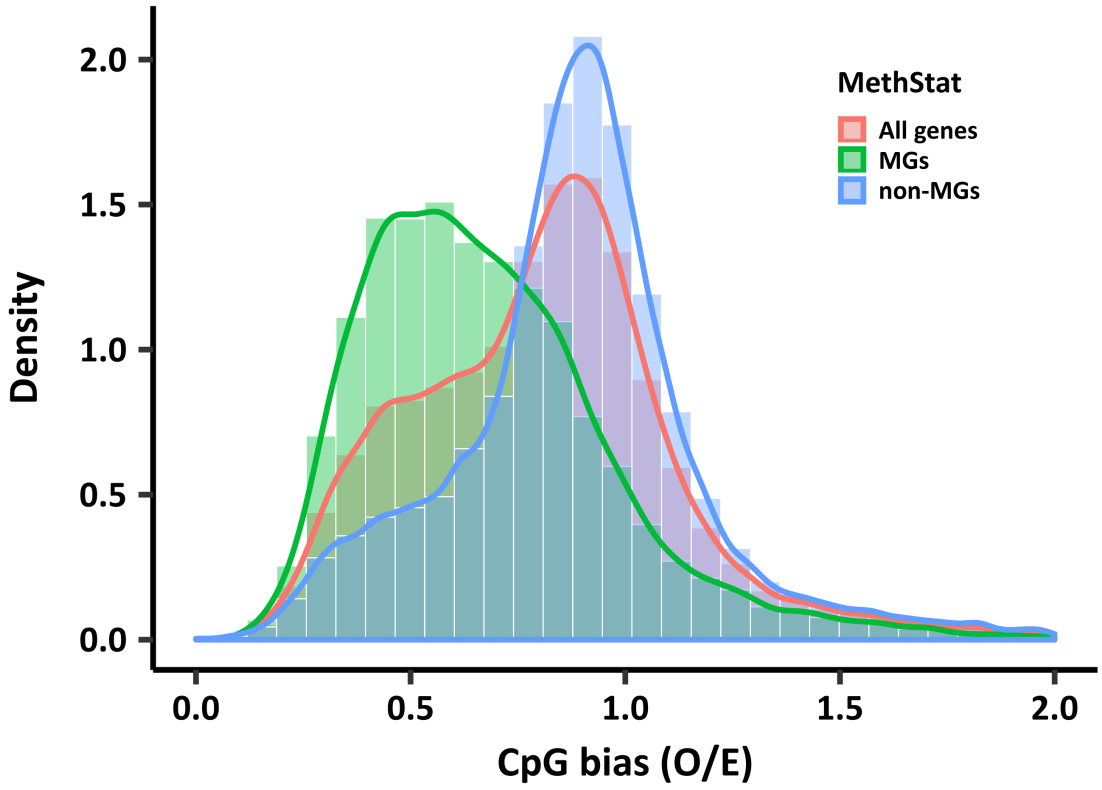
638 **Fig. S1. Circos visualization of different data at the genome-wide level**

639 (**a**) TE density. (**b**) Gene density. (**c**) Fraction of methylated CpGs in symbiotic treatment. (**d**)

640 Fraction of methylated CpGs in aposymbiotic treatment. (**e**) CpG content.

641

642



643

644 **Fig. S2. Methylated genes in Aiptasia have lower CpG O/E**

645 CpG distribution of methylated genes (represented by red curve) peaks at around 0.5, which is

646 lower than in unmethylated genes (represented by green curve) peaking at around 0.9. mC to T

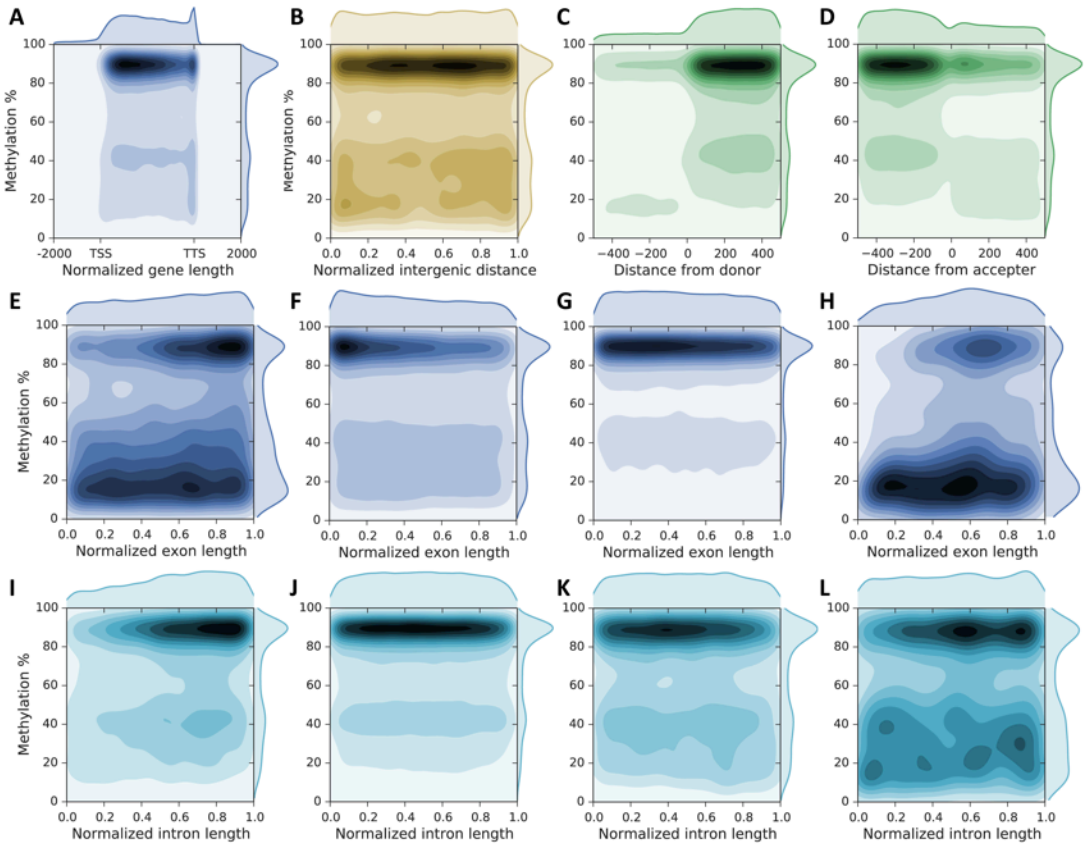
647 conversion skews the CpG O/E distribution of all genes as expected (represented by blue curve),

648 but methylated and unmethylated genes still show a large overlap of their CpG O/E distributions.

649 These results indicate that gene body methylation cannot be accurately inferred from CpG O/E in

650 Aiptasia.

651



652

653 **Fig. S3. Methylation patterns**

654 (**A**) DNA methylation is mainly located in the proximal part of gene bodies with slightly

655 decreasing levels towards the end. (**B**) Methylation pattern over intergenic regions. (**C**)

656 Methylation pattern around splice donor sites show increasing levels immediately after donor

657 sites. (**D**) Methylation pattern around acceptor sites show decreasing levels immediately after

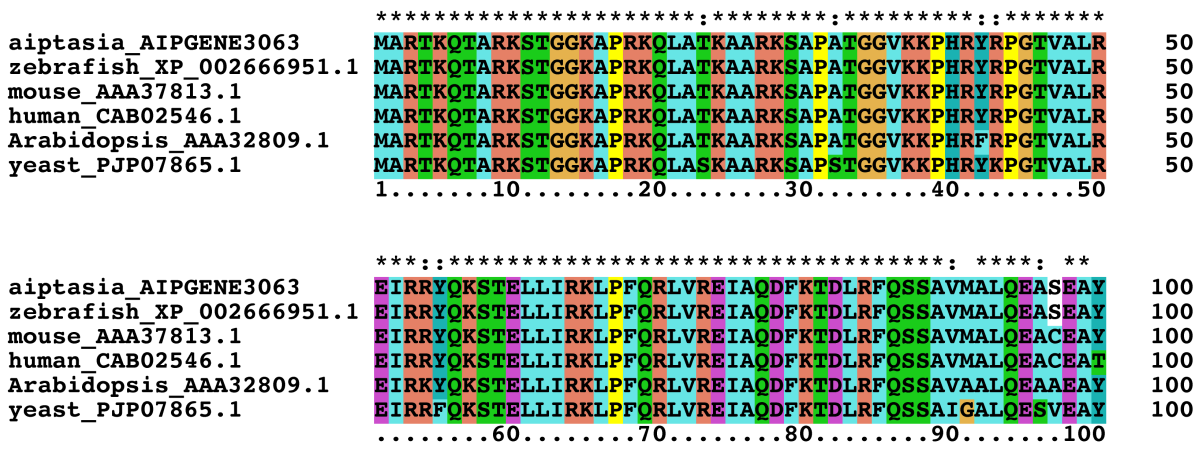
658 splice acceptor sites. (**E**) Methylation pattern over initial exons show increasing methylation

659 levels (3,147 exons with 35,885 methylation sites were used). (**F**) Methylation pattern over

660 internal exons show decreasing methylation levels (7,977 exons with 139,009 methylation sites

were used). (**G**) Methylation pattern over terminal exons show decreasing methylation levels (7,905 exons with 102,162 methylation sites were used). (**H**) Methylation pattern over introns from single-exon genes follow a similar trend as observed for multi exon genes with increasing methylation levels in the proximal and decreasing levels in the posterior part of the exon (298 exons with 4,735 methylation sites were used). (**I**) Methylation pattern over initial introns show increasing methylation levels (3,381 introns with 39,262 methylation sites were used). (**J**) Methylation pattern over internal introns maintain stable methylation levels (7,371 introns with 211,950 methylation sites were used). (**K**) Methylation levels over terminal introns decrease slightly (3,959 introns with 34,246 methylation sites were used). (**L**) Methylation levels over introns from one-intron genes change gently with initial increase followed by a decrease (1,055 introns with 10,709 methylation sites).

673



674

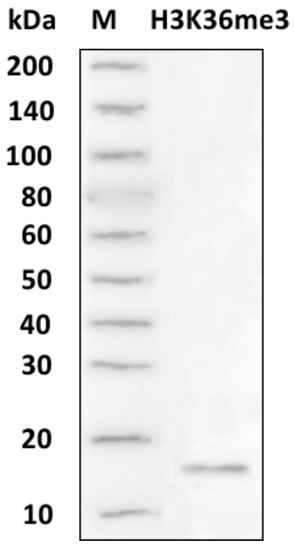
675 **Fig. S4. Sequence conservation of histone H3 homologs**

676 Sequence conservation of Aiptasia histone H3 protein and histone H3 homologs from species for
677 which antibody (ab9050, Abcam) has previously been validated. The N-terminal tail of Aiptasia
678 H3 is identical to the fragment from the zebrafish *Danio rerio* (the first 100 amino acid fragment
679 from human was used to produce this antibody).

680

681

682



683

684 **Fig. S5. Western blot**

685 Western blot result for antibody affinity validation, target band is 15kDa in size as expected from

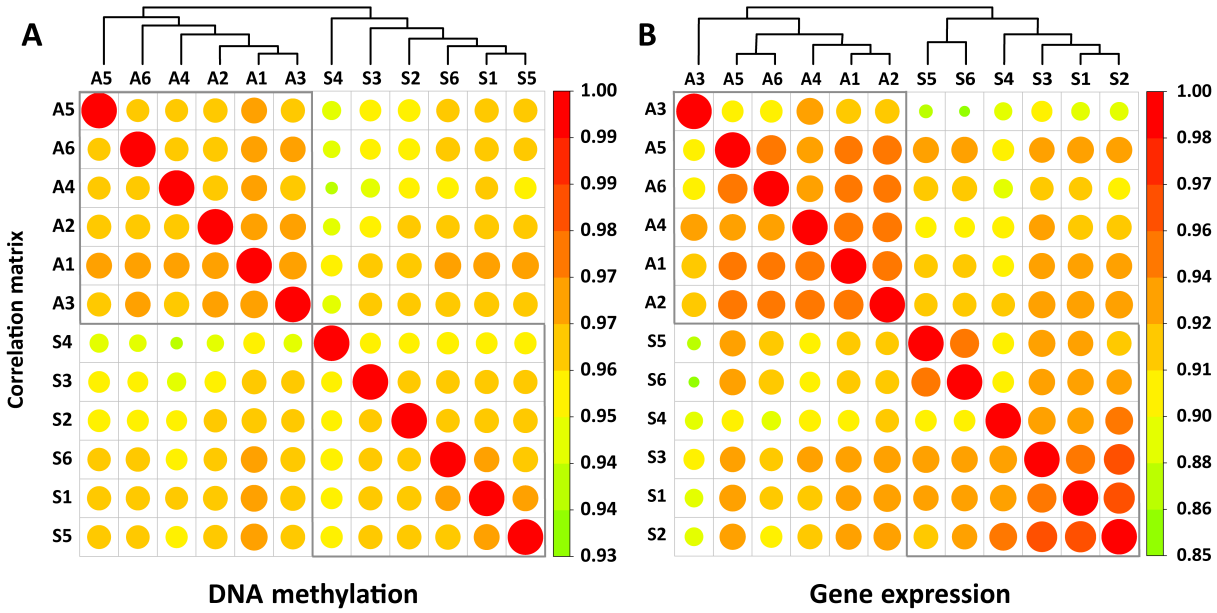
686 molecular weight analysis.

687

688

689

690



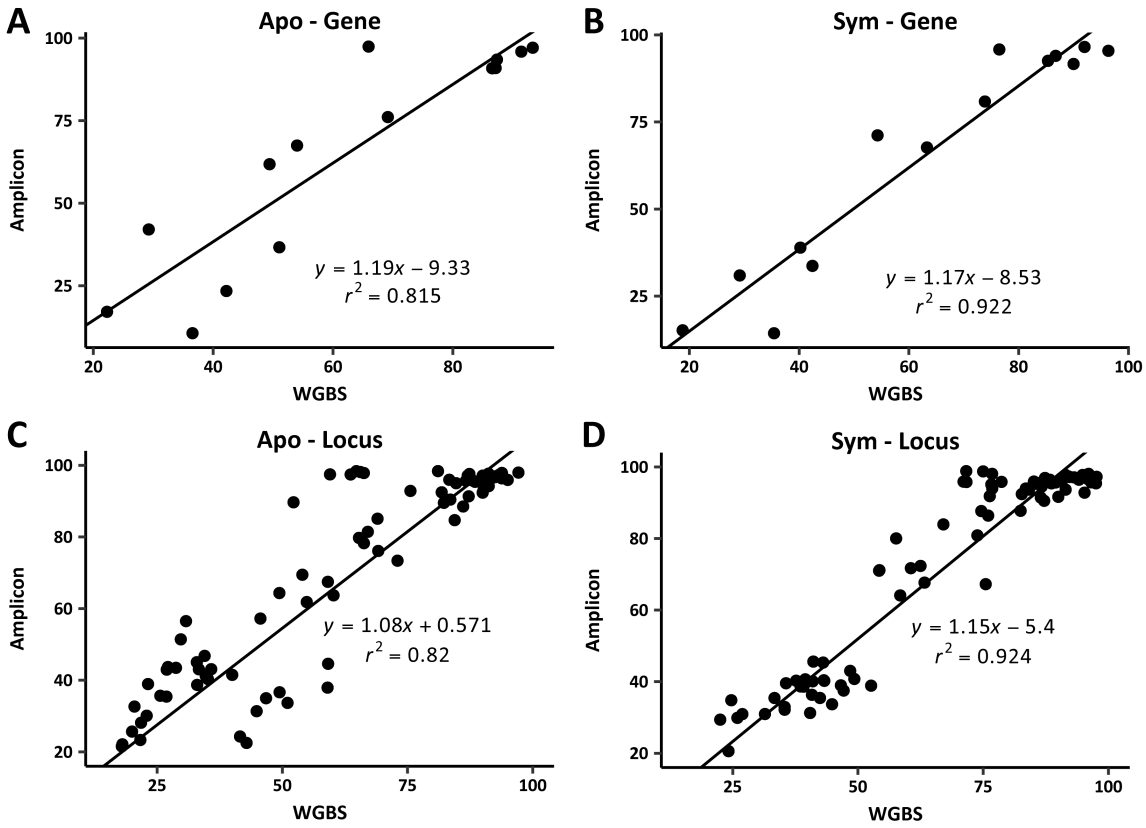
691

692 **Fig. S6. Correlation matrices of replicates**

693 Correlation matrices of replicates based on median DNA methylation level of genes (**A**) and log
694 gene expression values (base 2) (**B**). Replicates from the same treatments showed higher
695 correlation and clustered together both based on DNA methylation as well as gene expression
696 profiles, further supporting the findings obtained from the PCA analyses (figure 4) that changes
697 in DNA methylation and expression are treatment specific.

698

699



700

701 **Fig. S7. Validation of methylation level**

702 Validation of methylation level using bisulfite PCR on selected genes. (**A**, **B**) validation of
703 methylation level on genes (median methylation levels of methylated CpGs were used to
704 represent genes). (**C**, **D**) validation of methylation level on locus (methylated CpGs). WGBS:
705 whole genome bisulfite sequencing; Amplicon: MiSeq sequencing results of bisulfite PCR
706 amplicons on selected genes.

707

708

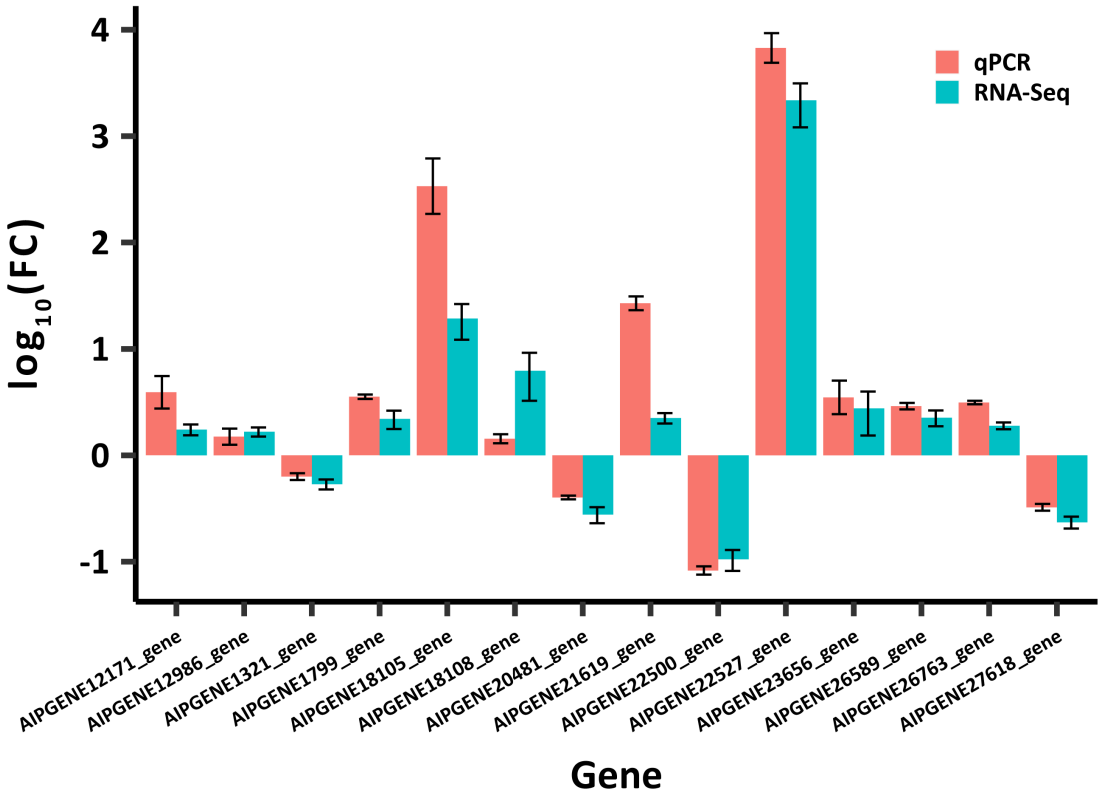


Fig. S8. qPCR validation of gene expression levels

Validation of gene expression changes using qPCR. Expression levels are shown as $\log_{10}$(fold change). All genes show similar expression changes as determined by RNA-seq and q-PCR.

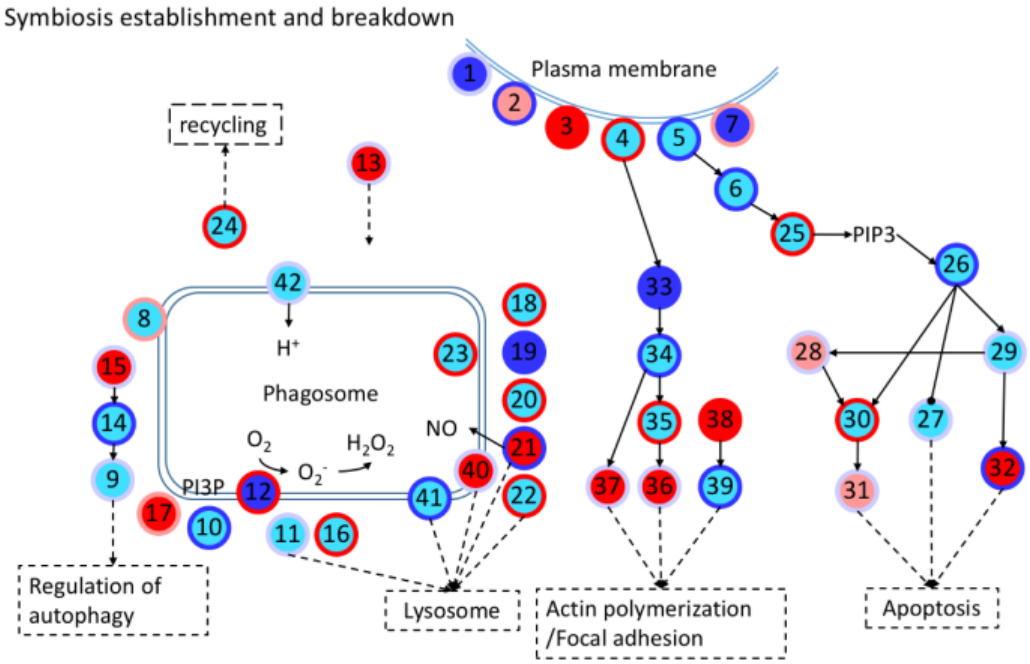


Fig. S9. Schematic diagram of symbiosis establishment and breakdown associated genes. Every node represents a category of genes, and generally has multiple corresponding genes. The inside colors of nodes represent the expression changes of corresponding genes, including non-DEGs (cyan), up-regulated (red), down-regulated (blue) and up- and down-regulated DEGs (light red). The colors of node edges represent the methylation level changes of corresponding genes, including non-DMGs (light blue), hypermethylated (red), hypomethylated (blue) and hyper- and hypo-methylated DMGs (light red). Numbers in circles denote genes/proteins as detailed below.

1. Complement receptor
2. Scavenger receptor
3. C-type lectin
4. Integrin
5. Toll-like receptor
6. Ras-related C3 botulinum toxin substrate 1 - rho family (RAC1)
7. Collagen
8. Vesicle-associated membrane protein (VAMP)
9. Autophagy-related protein 16 (ATG16)
10. Ras-related protein 5 (Rab5)
11. Ras-related protein 7 (Rab7)
12. NADPH oxidase (NOX)
13. Syntaxin 12
14. Autophagy-related protein 5 (ATG5)
15. Autophagy-related protein 10 (ATG10)
16. Programmed cell death 6-interacting protein
17. Sorting nexin (SNX)
18. Cytoplasmic dynein
19. Tubulin alpha chain (TUBA)
20. Tubulin beta chain (TUBB)
21. Nitric oxide synthase (NOS)
22. Lysosome-associated membrane glycoprotein/Cluster of differentiation (LAMP/CD)
23. Cathepsin L
24. Kinesin
25. Phosphatidylinositol 4,5-bisphosphate 3-kinase (PI3K)
26. RAC serine/threonine-protein kinase (AKT)
27. Bcl-2-antagonist of cell death (BAD)
28. TNF receptor-associated factor (TRAF)
29. Nuclear factor of kappa light polypeptide gene (NFKB)
30. Caspase 8 (CASP8)
31. Caspase 7 (CASP7)
32. Apoptosis regulator/Bcl-2 (BCL2)
33. Ras homolog (RHO)
34. Rho-associated protein kinase (ROCK)
35. Phosphatidylinositol 4-phosphate 5-kinase / Phosphatidylinositol 5-phosphate 4-kinase / Phosphatidylinositol 3-phosphate 5-kinase (PI4P5K/ PI5P4K/ PI3P5K)
36. Vinculin
37. Radixin
38. Profilin
39. Actin
40. CD63
41. Lysosomal-associated transmembrane protein
42. V-type proton ATPase

PI3P: phosphatidylinositol-3-phosphate

PIP3: Phosphatidylinositol (3,4,5)-trisphosphate

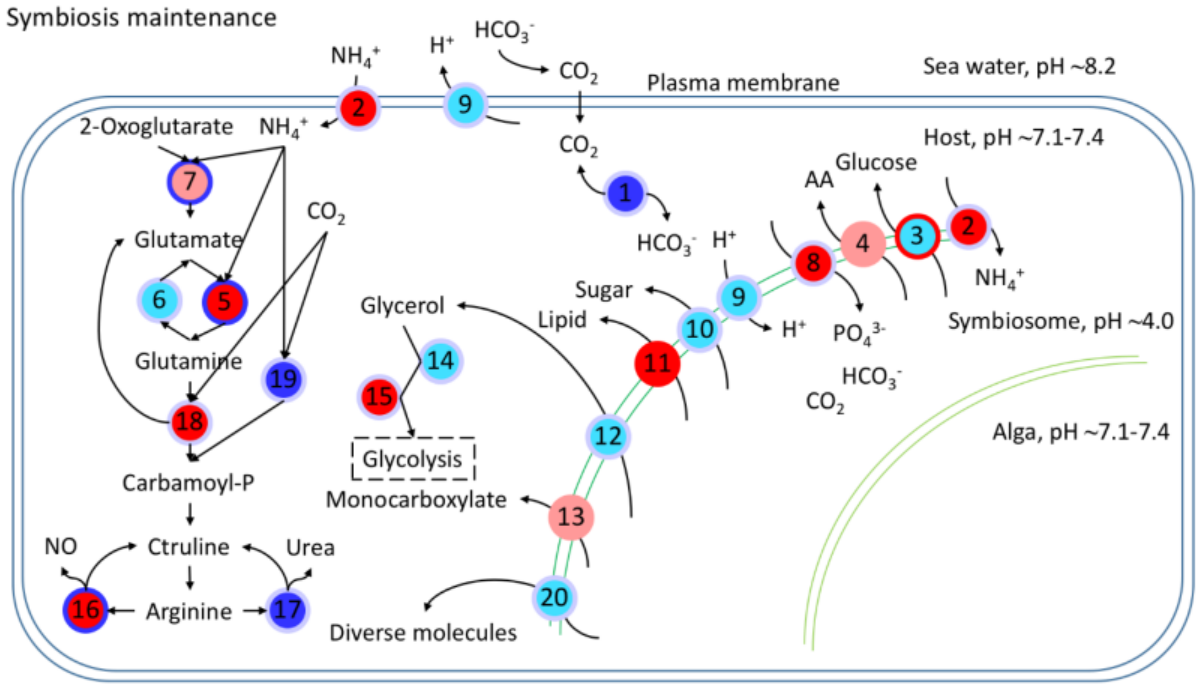


Fig. S10. Schematic diagram of symbiosis maintenance associated genes. Every node represents a category of genes, and generally has multiple corresponding genes. The inside colors of nodes represent the expression changes of corresponding genes, including non-DEGs (cyan), up-regulated (red), down-regulated (blue) and up- and down-regulated DEGs (light red). The colors of node edges represent the methylation level changes of corresponding genes, including non-DMGs (light blue), hypermethylated (red), hypomethylated (blue) and hyper- and hypo-methylated DMGs (light red).

1. Carbonic anhydrase (CA)
2. Ammonium transporter
3. Glucose transporter
4. Amino acid transporter
5. Glutamine synthetase (GS)
6. Glutamate synthase
7. Glutamate dehydrogenase (GDH)
8. Phosphate transporter
9. V-type proton ATPase
10. Sugar transporter
11. Lipid transfer protein
12. Aquaporin 3 (Glycerol transporter)
13. Monocarboxylate transporter
14. Alcohol dehydrogenase
15. Aldehyde dehydrogenase
16. Nitric oxide synthase (NOS)
17. Arginase
18. Carbamoyl-phosphate synthase / Aspartate carbamoyltransferase / Dihydroorotase (CAD)
19. Carbamoyl-phosphate synthase (ammonia) (CPS1)
20. ABC transporter

725

726

References

Adrian Alexa JR. 2016. topGO: Enrichment Analysis for Gene Ontology. *R package version 2240*.

Anitei M, Stange C, Parshina I, Baust T, Schenck A, Raposo G, Kirchhausen T, Hoflack B. 2010. Protein complexes containing CYFIP/Sra/PIR121 coordinate Arf1 and Rac1 signalling during clathrin-AP-1-coated carrier biogenesis at the TGN. *Nat Cell Biol* **12**: 330-340.

Aranda M, Li Y, Liew YJ, Baumgarten S, Simakov O, Wilson MC, Piel J, Ashoor H, Bougouffa S, Bajic VB et al. 2016. Genomes of coral dinoflagellate symbionts highlight evolutionary adaptations conducive to a symbiotic lifestyle. *Scientific Reports* **6**: 39734.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT et al. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**: 25-29.

Baietti MF, Zhang Z, Mortier E, Melchior A, Degeest G, Geeraerts A, Ivarsson Y, Depoortere F, Coomans C, Vermeiren E et al. 2012. Syndecan-syntenin-ALIX regulates the biogenesis of exosomes. *Nat Cell Biol* **14**: 677-685.

Baumgarten S, Simakov O, Esherick LY, Liew YJ, Lehnert EM, Michell CT, Li Y, Hambleton EA, Guse A, Oates ME et al. 2015. The genome of Aiptasia, a sea anemone model for coral symbiosis. *Proc Natl Acad Sci U S A* **112**: 11893-11898.

Bird A. 2002. DNA methylation patterns and epigenetic memory. *Genes Dev* **16**: 6-21.

Chen MC, Cheng YM, Hong MC, Fang LS. 2004. Molecular cloning of Rab5 (ApRab5) in Aiptasia pulchella and its retention in phagosomes harboring live zooxanthellae. *Biochem Biophys Res Commun* **324**: 1024-1033.

Chen MC, Cheng YM, Sung PJ, Kuo CE, Fang LS. 2003. Molecular identification of Rab7 (ApRab7) in Aiptasia pulchella and its exclusion from phagosomes harboring zooxanthellae. *Biochem Biophys Res Commun* **308**: 586-595.

Cook CB, Mullerparker G, Delia CF. 1992. Ammonium Enhancement of Dark Carbon Fixation and Nitrogen Limitation in Symbiotic Zooxanthellae - Effects of Feeding and Starvation of the Sea Anemone Aiptasia Pallida. *Limnology and Oceanography* **37**: 131-139.

Davidson MG. 2002. Protecting coral reefs: The principal national and international legal instruments. *The Harvard Environmental Law Review* **26**: 48.

Davy SK, Allemand D, Weis VM. 2012. Cell biology of cnidarian-dinoflagellate symbiosis. *Microbiology and molecular biology reviews : MMBR* **76**: 229-261.

Dixon GB, Bay LK, Matz MV. 2016. Evolutionary Consequences of DNA Methylation in a Basal Metazoan. *Mol Biol Evol* **33**: 2285-2293.

Dixon GB, Bay LK, Matz MV. 2017. Patterns of gene body methylation predict coral fitness in new environments. *bioRxiv* doi:10.1101/184457.

Downs CA, Kramarsky-Winter E, Martinez J, Kushmaro A, Woodley CM, Loya Y, Ostrander GK. 2009. Symbiophagy as a cellular mechanism for coral bleaching. *Autophagy* **5 2**: 211-216.

Dunn SR, Schnitzler CE, Weis VM. 2007. Apoptosis and autophagy as mechanisms of dinoflagellate symbiont release during cnidarian bleaching: every which way you lose. *Proceedings of the Royal Society B: Biological Sciences* **274**: 3079-3085.

Feng S, Cokus SJ, Zhang X, Chen PY, Bostick M, Goll MG, Hetzel J, Jain J, Strauss SH, Halpern ME et al. 2010. Conservation and divergence of methylation patterning in plants and animals. *Proc Natl Acad Sci U S A* **107**: 8689-8694.

Foret S, Kucharski R, Pellegrini M, Feng S, Jacobsen SE, Robinson GE, Maleszka R. 2012. DNA methylation dynamics, metabolic fluxes, gene splicing, and alternative phenotypes in honey bees. *Proc Natl Acad Sci U S A* **109**: 4968-4973.

Gavery MR, Roberts SB. 2010. DNA methylation patterns provide insight into epigenetic regulation in the Pacific oyster (Crassostrea gigas). *BMC Genomics* **11**: 483.

Gonzalez-Romero R, Suarez-Ulloa V, Rodriguez-Casariego J, Garcia-Souto D, Diaz G, Smith A, Pasantes JJ, Rand G, Eirin-Lopez JM. 2017. Effects of Florida Red Tides on histone variant expression and DNA methylation in the Eastern oyster Crassostrea virginica. *Aquatic Toxicology* **186**: 196-204.

778 Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng
779 Q et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat*
780 *Biotechnol* **29**: 644-652.
781 Grover R, Maguer J-F, Allemand D, Ferrier-Pagès C. 2008. Uptake of dissolved free amino acids by the
782 scleractinian coral *Stylophora pistillata*. *Journal of Experimental Biology* **211**: 860-865.
783 Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M et
784 al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for
785 reference generation and analysis. *Nat Protoc* **8**: 1494-1512.
786 Hastie TJ, Pregibon D. 1992. Generalized linear models. In *Statistical Models in S*.
787 He X-J, Chen T, Zhu J-K. 2011. Regulation and function of DNA methylation in plants and animals. *Cell Res* **21**:
788 442-465.
789 Janke R, Iavarone AT, Rine J. 2017. Oncometabolite D-2-Hydroxyglutarate enhances gene silencing through
790 inhibition of specific H3K36 histone demethylases. *eLife* **6**: e22451.
791 Jones PA. 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* **13**:
792 484-492.
793 Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**: 27-30.
794 Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2016. KEGG as a reference resource for gene and
795 protein annotation. *Nucleic Acids Res* **44**: D457-462.
796 Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nat Meth*
797 **12**: 357-360.
798 Krueger F, Andrews SR. 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications.
799 *Bioinformatics* **27**: 1571-1572.
800 Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an
801 information aesthetic for comparative genomics. *Genome Res* **19**: 1639-1645.
802 Lämke J, Bäurle I. 2017. Epigenetic and chromatin-based mechanisms in environmental stress adaptation and stress
803 memory in plants. *Genome Biology* **18**: 124.
804 Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Meth* **9**: 357-359.
805 Latysheva N, Muratov G, Rajesh S, Padgett M, Hotchin NA, Overduin M, Berditchevski F. 2006. Syntenin-1 is a
806 new component of tetraspanin-enriched microdomains: mechanisms and consequences of the interaction of
807 syntenin-1 with CD63. *Mol Cell Biol* **26**: 7707-7718.
808 Lehnert EM, Mouchka ME, Burriesci MS, Gallo ND, Schwarz JA, Pringle JR. 2014. Extensive Differences in Gene
809 Expression Between Symbiotic and Aposymbiotic Cnidarians. *G3: Genes|Genomes|Genetics* **4**: 277-295.
810 Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference
811 genome. *BMC Bioinformatics* **12**: 1-16.
812 Liew YJ, Zoccola D, Li Y, Tambutté E, Venn AA, Michell CT, Cui G, Deutekom ES, Kaandorp JA, Voolstra CR et
813 al. 2017. Epigenome-associated phenotypic acclimatization to ocean acidification in a reef-building coral.
814 *bioRxiv* doi:10.1101/188227.
815 Lin S, Cheng S, Song B, Zhong X, Lin X, Li W, Li L, Zhang Y, Zhang H, Ji Z et al. 2015. The Symbiodinium
816 kawagutii genome illuminates dinoflagellate gene expression and coral symbiosis. *Science* **350**: 691-694.
817 Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011* **17**.
818 McAuley PJ, Smith DC. 1982. The Green Hydra Symbiosis. V. Stages in the Intracellular Recognition of Algal
819 Symbionts by Digestive Cells. *Proceedings of the Royal Society B: Biological Sciences* **216**: 7-23.
820 Meyer E, Weis VM. 2012. Study of cnidarian-algal symbiosis in the "omics" age. *The Biological bulletin* **223**: 44-
821 65.
822 Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. 2007. KAAS: an automatic genome annotation and
823 pathway reconstruction server. *Nucleic Acids Res* **35**: W182-185.
824 Muscatine L. 1990. *The role of symbiotic algae in carbon and energy flux in reef corals*. Elsevier Science Publishing
825 Company.
826 Nanty L, Carbajosa G, Heap GA, Ratnieks F, van Heel DA, Down TA, Rakyan VK. 2011. Comparative
827 methylomics reveals gene-body H3K36me3 in Drosophila predicts DNA methylation and CpG landscapes
828 in other invertebrates. *Genome Research* **21**: 1841-1850.
829 Neri F, Rapelli S, Krepelova A, Incarnato D, Parlato C, Basile G, Maldotti M, Anselmi F, Oliviero S. 2017.
830 Intragenic DNA methylation prevents spurious transcription initiation. *Nature* **543**: 72-77.
831 Neubauer EF, Poole AZ, Detournay O, Weis VM, Davy SK. 2016. The scavenger receptor repertoire in six
832 cnidarian species and its putative role in cnidarian-dinoflagellate symbiosis. *PeerJ* **4**: e2692.

833 Oakley CA, Ameismeier MF, Peng L, Weis VM, Grossman AR, Davy SK. 2016. Symbiosis induces widespread
834 changes in the proteome of the model cnidarian Aiptasia. *Cellular Microbiology* **18**: 1009-1023.
835 Pertea M, Pertea GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. 2015. StringTie enables improved
836 reconstruction of a transcriptome from RNA-seq reads. *Nat Biotech* **33**: 290-295.
837 Putnam HM, Davidson JM, Gates RD. 2016. Ocean acidification influences host DNA methylation and phenotypic
838 plasticity in environmentally susceptible corals. *Evolutionary Applications* **9**: 1165-1178.
839 R Core Team. 2016. R: A Language and Environment for Statistical Computing. *https://wwwR-projectorg/*.
840 Radecker N, Pogoreutz C, Voolstra CR, Wiedenmann J, Wild C. 2015. Nitrogen cycling in corals: the key to
841 understanding holobiont functioning? *Trends Microbiol* **23**: 490-497.
842 Rädecker N, Pogoreutz C, Wild C, Voolstra CR. 2017. Stimulated Respiration and Net Photosynthesis in Cassiopeia
843 sp. during Glucose Enrichment Suggests in hospite CO2 Limitation of Algal Endosymbionts. *Frontiers in*
844 *Marine Science* **4**.
845 Rando OJ, Verstrepen KJ. 2007. Timescales of Genetic and Epigenetic Inheritance. *Cell* **128**: 655-668.
846 Richards EJ. 2006. Inherited epigenetic variation - revisiting soft inheritance. *Nat Rev Genet* **7**: 395-401.
847 Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis
848 of digital gene expression data. *Bioinformatics* **26**: 139-140.
849 Rodriguez-Lanetty M, Phillips WS, Weis VM. 2006. Transcriptome analysis of a cnidarian-dinoflagellate mutualism
850 reveals complex modulation of host gene expression. *BMC Genomics* **7**: 23.
851 Soboleva TA, Parker BJ, Nekrasov M, Hart-Smith G, Tay YJ, Tng W-Q, Wilkins M, Ryan D, Tremethick DJ. 2017.
852 A new link between transcriptional initiation and pre-mRNA splicing: The RNA binding histone variant
853 H2A.B. *PLOS Genetics* **13**: e1006633.
854 Spalding MD, Grenfell AM. 1997. New estimates of global and regional coral reef areas. *Coral Reefs* **16**: 225-230.
855 Sunagawa S, Wilson EC, Thaler M, Smith ML, Caruso C, Pringle JR, Weis VM, Medina M, Schwarz JA. 2009.
856 Generation and analysis of transcriptomic resources for a model system on the rise: the sea anemone
857 Aiptasia pallida and its dinoflagellate endosymbiont. *BMC Genomics* **10**: 258.
858 Suzuki MM, Bird A. 2008. DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet* **9**:
859 465-476.
860 Suzuki MM, Kerr AR, De Sousa D, Bird A. 2007. CpG methylation is targeted to transcription units in an
861 invertebrate genome. *Genome Res* **17**: 625-631.
862 Sylvain JC. 2006. How to Protect A Coral Reef: The Public Trust Doctrine and the Law of the Sea. *Sustainable*
863 *Development Law & Policy* **7**: 4.
864 Torda G, Donelson JM, Aranda M, Barshis DJ, Bay L, Berumen ML, Bourne DG, Cantin N, Foret S, Matz M et al.
865 2017. Rapid adaptive responses to climate change in corals. *Nature Clim Change* **7**: 627-636.
866 Tucker KL. 2001. Methylated cytosine and the brain: a new base for neuroscience. *Neuron* **30**: 649-652.
867 Vastenhouw NL, Zhang Y, Woods IG, Imam F, Regev A, Liu XS, Rinn J, Schier AF. 2010. Chromatin signature of
868 embryonic pluripotency is established during genome activation. *Nature* **464**: 922.
869 Venables WN, Ripley BD. 2002. Exploratory Multivariate Analysis. In *Modern Applied Statistics with S*,
870 doi:10.1007/978-0-387-21706-2. Springer-Verlag New York.
871 Voolstra CR. 2013. A journey into the wild of the cnidarian model system Aiptasia and its symbionts. *Molecular*
872 *Ecology* **22**: 4366-4368.
873 Wang J, Douglas A. 1998a. Nitrogen recycling or nitrogen conservation in an alga-invertebrate symbiosis? *The*
874 *Journal of Experimental Biology* **201**: 2445-2453.
875 Wang J, Douglas AE. 1998b. Nitrogen recycling or nitrogen conservation in an alga-invertebrate symbiosis? *J Exp*
876 *Biol* **201 (Pt 16)**: 2445-2453.
877 Wang X, Wheeler D, Avery A, Rago A, Choi J-H, Colbourne JK, Clark AG, Werren JH. 2013. Function and
878 Evolution of DNA Methylation in *Nasonia vitripennis*. *PLoS Genet* **9**: e1003872.
879 Wang Y, Leung FC. 2008. GC content increased at CpG flanking positions of fish genes compared with sea squirt
880 orthologs as a mechanism for reducing impact of DNA methylation. *PLoS One* **3**: e3612.
881 Wollmann H, Stroud H, Yelagandula R, Tarutani Y, Jiang D, Jing L, Jamge B, Takeuchi H, Holec S, Nie X et al.
882 2017. The histone H3 variant H3.3 regulates gene body DNA methylation in Arabidopsis thaliana. *Genome*
883 *Biology* **18**: 94.
884 Wu S-F, Zhang H, Cairns BR. 2011. Genes for embryo development are packaged in blocks of multivalent
885 chromatin in zebrafish sperm. *Genome Research* **21**: 578-589.
886 Xiang T, Hambleton EA, DeNofrio JC, Pringle JR, Grossman AR. 2013. Isolation of clonal axenic strains of the
887 symbiotic dinoflagellate Symbiodinium and their growth and host specificity(1). *J Phycol* **49**: 447-458.

888 Zeng J, Yi SV. 2010. DNA Methylation and Genome Evolution in Honeybee: Gene Length, Expression, Functional
889 Enrichment Covary with the Evolutionary Signature of DNA Methylation. *Genome Biology and Evolution*
890 **2**: 770-780.
891 Zilberman D. 2017. An evolutionary case for functional gene body methylation in plants and animals. *Genome*
892 *Biology* **18**: 87.
893