Article - Resources

Genome-wide disruption of DNA methylation by 5-aza-2'-deoxycytidine in a parasitoid wasp

Nicola Cook[1¶*], Darren J Parker[2,3¶], Frances Turner[4], Eran Tauber[5], Bart A Pannebakker[6] and David M Shuker[1]

1. School of Biology, University of St Andrews, UK

2. Department of Ecology and Evolution, University of Lausanne, Switzerland

3. Swiss Institute of Bioinformatics, Lausanne, Switzerland

4. Edinburgh Genomics, School of Biological Sciences, University of Edinburgh, UK

5. Faculty of Natural Sciences, University of Haifa, Israel

6. Laboratory of Genetics, Wageningen University, The Netherlands

*Corresponding author

E-mail: nc52@st-andrews.ac.uk

¶These authors contributed equally

Abstract

DNA methylation of cytosine residues across the genome influences how many genes and phenotypes are regulated. As such, understanding the role of DNA methylation and other epigenetic mechanisms has become very much a part of mapping genotype to phenotype, a major question in evolutionary biology. Ideally, we would like to manipulate DNA methylation patterns on a genome-wide scale, to elucidate the role of epigenetic modifications in phenotypic expression. Recently, the demethylating agent 5-aza-2'-deoxycytidine (5-aza-dC; commonly used in the epigenetic treatment of certain cancers), has been deployed to explore the epigenetic regulation of a number of traits of interest to evolutionary ecologists. Recently, we showed that treatment with 5-aza-dC shifted patterns of sex allocation as predicted by genomic conflict theory in the parasitoid wasp *Nasonia vitripennis*. This was the first (albeit indirect) experimental evidence for genomic conflict over sex allocation facilitated by DNA methylation. However, this work lacked confirmation of the effects of 5-aza-dC on DNA methylation, drawing commentary on the efficacy of 5-aza-dC in a novel system. Here, using whole-genome bisulphite sequencing, we demonstrate unequivocally that 5-aza-dC disrupts methylation across the *Nasonia vitripennis* genome. We show that disruption leads to both hypo- and hyper-methylation, may vary across tissues and time of sampling, and that the effects of 5-aza-dC are context- and sequence specific. We conclude that 5-aza-dC has the potential to be repurposed as a tool in evolutionary ecology for studying the role of DNA methylation.

52 **Introduction**

53 The genomics revolution has transformed our understanding of how genomes build phenotypes. The
54 extraordinary diversity of morphologies, behaviours, and physiological adaptations we see across the
55 natural world are not necessarily the product of different sets of genes, or different numbers of genes
56 in organisms, but rather differences in how, where and when genes are expressed during
57 development. The genomics revolution has therefore been as much about understanding gene
58 regulation as it has been about identifying allelic variants at genetic loci (Biémont 2010, Ashbrook *et*
59 *al.* 2018). At the forefront of our growing knowledge of gene regulation has been the role of epigenetic
60 modification of DNA, including DNA methylation. In the vast majority of cases, DNA methylation occurs
61 at the C5 position of the cytosine ring in CpG dinucleotides and is catalysed by DNA methyltransferase
62 enzymes (Dnmts). Three types of Dnmt are known in eukaryotes; Dnmt3 establishes DNA methylation
63 patterns *de novo*, Dnmt1 maintains these patterns, and Dnmt2 is involved in tRNA methylation (Lyko
64 2018). In mammals, broadly, DNA methylation is found throughout the genome (except at CpG islands
65 near the promoters of genes) and is associated with transcriptional repression (Feng *et al.* 2010,
66 Zemach *et al.* 2010). By contrast, in insects, DNA methylation is concentrated in gene bodies and is
67 associated with more stable patterns of gene expression (Bewick *et al.* 2017). Notably, vertebrate and
68 invertebrate DNA methylation varies greatly in its extent throughout the genome, with the proportion
69 of methylated CpGs much lower in invertebrates: 0-14% as compared to 60-90% in mammals (Bewick
70 *et al.* 2017, Glastad *et al.* 2011). The evolutionary reasons for these differences are not yet clear,
71 although it is thought that gene body methylation is an ancient feature of eukaryote genomes (Zemach
72 *et al.* 2010). However, DNA methylation status influences how genes (and phenotypes) are expressed
73 in many organisms, including reproductive vs non-reproductive worker bumblebees (Amarasinghe *et*
74 *al.* 2014), caste differentiation in honey bees (Kucharski *et al.* 2008, Lyko *et al.* 2010),
75 transgenerational response to the presence of predators in *Daphnia magna* (Schield *et al.* 2016), and
76 between lateral plate morphotypes in Threespine Stickleback (Smith *et al.* 2014). Therefore, DNA
77 methylation patterns are very much part of mapping genotypes to phenotypes.

78 The flexible control of gene expression provided by mechanisms such as DNA methylation has
79 also caught the interest of biologists interested in the underlying mechanisms of phenotypic plasticity,
80 and how such plasticity may evolve (Bonduriansky and Day 2009, 2018). Moreover, there has also
81 been great interest in the extent to which epigenetic modifications may be inherited across
82 generations (Bonduriansky and Day 2009, Burggren 2016, Lind and Spagopoulou 2018). Transmission
83 of epigenetic state allows transmission of parent-of-origin information, or so-called genomic
84 imprinting, setting the scene for genomic conflict (Haig 2002, Patten *et al.* 2014). Genomic conflict
85 arises when maternally- and paternally-inherited alleles have different evolutionary optima for traits

86 expressed in an individual (Haig 1997, 2002). A classic example is the expression of maternally- or
87 paternally-inherited alleles during embryonic development in mammals: paternally-inherited alleles
88 are selected to extract more resources from the mother *in utero* compared to maternally-inherited
89 alleles (Moore and Haig 1991). This conflict is now known to be, at least in part, facilitated by parent-
90 of-origin specific gene expression of insulin-like growth factor II (Igf2) and Igf2r. In mammals, the Igf2
91 gene promotes growth and cellular differentiation during development, and also regulates the
92 placental supply of nutrients and the demand of nutrients by the foetus (Constancia *et al.* 2002).

93 To further understand the role epigenetic modifications such as DNA methylation have in
94 influencing phenotypes and their evolution, we need robust ways of manipulating them. Here we
95 consider the efficacy of the chemical 5-aza-2'-deoxycytidine (5-aza-dC) as a way of manipulating DNA
96 methylation, using the parasitoid wasp *Nasonia vitripennis*. The extent and nature of DNA methylation
97 across the *N. vitripennis* genome has recently been characterised (Wang *et al.* 2013, Pegoraro *et al.*
98 2016), and the presence in the genome of a full "methylation toolkit" has also been confirmed (Werren
99 *et al.* 2010). 5-aza-dC is well-known from the cancer literature as a hypomethylating agent and is used
100 in the treatment of acute myeloid leukaemia and myelodysplastic syndrome (Bryan *et al.* 2011,
101 Momparler 2012, Seelan *et al.* 2018). In cancer patients, many genes that suppress leukemogenesis in
102 humans are silenced by aberrant DNA methylation. 5-aza-dC acts as a demethylating agent to
103 reactivate these genes (see below for details of action). In recent years, 5-aza-dC has been used across
104 a range of species as an experimental tool to explore how DNA methylation influences phenotypes
105 including diapause in *Nasonia vitripennis* (Pegoraro *et al.* 2016) and worker reproductive status in
106 bumble bees (Amarasinghe *et al.* 2014). Previously, we used 5-aza-dC to explore whether DNA
107 methylation was involved in the control of facultative sex allocation in *N. vitripennis* (Cook *et al.* 2015).

108 Like all Hymenoptera, *N. vitripennis* is haplodiploid, with males arising from unfertilised
109 haploid eggs and females arising from fertilised diploid eggs. Females have putative control over the
110 sex of their offspring, releasing sperm to produce daughters or not to produce sons. Female *Nasonia*
111 allocate sex broadly in line with the predictions of Local Mate Competition theory (Hamilton 1967,
112 Taylor and Bulmer 1980, Werren 1980, Werren 1983, Shuker and West 2004, Burton-Chellew *et al.*
113 2008; reviewed by West 2009). Females laying eggs alone on their blowfly pupae hosts produce very
114 female-biased offspring sex ratios to minimise competition amongst her sons for mating
115 opportunities. As more females (foundresses) lay eggs on the same host(s) together, females produce
116 less female-biased sex ratios, as LMC is reduced and the fitness gained through the production of sons
117 and daughters moves towards equality (Hamilton 1967). When treated with 5-aza-dC, the facultative
118 sex allocation response in female *Nasonia* was maintained, suggesting that this plastic behaviour in
119 and of itself did not require certain patterns of methylation. Rather, the sex ratios (as proportion male)

120 produced by females were shifted slightly upwards, with more sons produced by treated females than
121 by controls. This subtle shift is as predicted by genomic conflict over sex allocation theory, which
122 requires some form of genomic imprinting (Wild and West 2008). Assuming that DNA methylation is
123 the mechanism by which parent-of-origin allelic information is carried in *Nasonia* (although see Wang
124 *et al.* 2016), our data therefore provided the first (albeit indirect) evidence for genomic conflict
125 influencing sex allocation (Cook *et al.* 2015a).

126 Recently, Ellers *et al.* (2019) challenged the effects of 5-aza-dC as a general modifier of
127 methylation status, considering CpGs in nine genes in *Nasonia vitripennis*. Ellers *et al.* (2009) failed to
128 find changes in methylation status across any of the 155 CpGs in their study, questioning the action of
129 5-aza-dC in *Nasonia*, and hence of course the interpretation of the effects of 5-aza-dC in our earlier
130 study (Cook *et al.* 2015a). Fully negative results are of course very hard to interpret, but Ellers *et al.*
131 (2009) nonetheless raised valid concerns of the efficacy of the chemical. As noted above, levels of
132 methylation in insects are generally low, and a genome-wide study would provide greater resolution
133 to explore the effects of this chemical in an ecologically valuable model system. Here we consider the
134 effects of 5-aza-dC across the whole genome using a Bisulfite sequencing (BS-Seq) experiment. There
135 is evidence to suggest that the effects of 5-aza-dC are unlikely to be random across methylated CpGs
136 in a given genome. For instance, in mammals, genomic context can influence the extent to which 5-
137 aza-dC is effective as a demethylating agent. DNA sequence context, the distribution of transcription
138 factor binding sites, and chromatin structure are known to influence the DNA methylation patterns
139 produced by 5-aza-dC exposure (Mossman *et al.* 2010, Hagemann *et al.* 2011, Ramos *et al.* 2015).
140 Although patterns of methylation differ among eukaryotes, the epigenetic modification itself is
141 chemically identical. Here, we verify that 5-aza-dC influences patterns of methylation across the
142 genome in *N. vitripennis*. Altered methylation patterns produced by 5-aza-dC exposure were non-
143 random, with enrichment in genes associated with transcription-factor activity and sequence specific
144 binding.

145

146 **Results**

147 Treatment with 5-aza-dC had wide-ranging effects on DNA methylation across the genome. First,
148 samples clustered strongly by 5-aza-dC exposure regime at the gene-level in both t-SNE visualisation
149 (Fig 1) and in principle component analyses (Fig 2), indicating that 5-aza-dC has a strong effect on the
150 methylation of genes. Both analyses show strong clustering, and that wasps exposed to 5-aza-dC for
151 24 hours are more like controls than wasps exposed to 5-aza-dC for 48 hours. In addition, the PCA
152 analysis showed clear separation along PCA1 (which presumably is heavily associated with exposure

153 regime; Fig 2). Importantly though, this analysis also emphasises the extent of variation in response
154 across the three treatment regimes, which is smallest for the controls, and largest for the 48 hour-
155 exposed wasps (note the spread along PCA axis 2 in Fig 2). In addition, both cluster analyses placed
156 the controls as some way intermediate between the two 5-aza-dC exposure treatments, which we
157 consider further in the next two paragraphs. Both analyses also show that exposure regime dominated
158 the clustering, over and above of any effects of either tissue-type (wasp heads vs bodies) or collection
159 time (hours post-exposure), suggesting that this is the key effect. When we consider clustering at the
160 level of individual CpG loci, clustering is much less obvious, using both t-SNE and PCA. However, the
161 control treatment appears to cluster within the others in the t-SNE analyses, and the 24-hour and 48-
162 hour treatments do separate to some extent, e.g. with the 24h samples more to the top and right of
163 Fig S1 and the 48h samples more to the bottom and left (see Figs S1 and S2). The differences between
164 the individual CpG and gene-level cluster analyses are likely to be due to a combination of low
165 coverage and a greater stochasticity of methylation at the individual CpG level (see discussion).

166 Second, a more in-depth examination of the data using a generalised linear modelling (GLM)
167 approach to examine differential methylation at the gene level found that 5-aza-dC exposure had a
168 significant effect on the methylation status of the majority of genes examined; exposure treatment
169 was significant as a main effect at an FDR level of 5% for 8,725 out of 11,582 genes, or 75.3% of genes
170 (see also Fig S3). Moreover, both tissue-type and collection time also influenced methylation status
171 across a large number of genes (significant as main effects in the GLMs for 3833 and 3566 genes
172 respectively). Significant interactions between factors were also present for the majority of genes
173 (exposure regime * collection time = 10,776 genes, exposure regime * tissue type = 10,484 genes,
174 collection time * tissue type = 5,364 genes, and exposure * collection time * tissue type = 10,484
175 genes), showing widespread evidence for significant interplay between these factors. These results
176 emphasise the context-dependent nature of the action of 5-aza-dC. Full results from the GLMs are
177 presented in Table S1.

178 Interestingly, it appears that methylation tends to increase initially in response to exposure to
179 5-aza-dC, an effect that is noticeable after 24h of exposure, before decreasing after 48h of exposure
180 (Fig 3A). This means that the length of time individuals are exposed to the chemical may well influence
181 what kind of methylation changes are observed; in other words, both hyper- and hypo-methylation
182 can result from 5-aza-dC treatment. In terms of when wasps were collected after exposure, 3,566
183 genes displayed significantly altered methylation in association with collection time, with a more
184 pronounced decrease in methylation in samples collected 48h post-exposure relative to 24h post-
185 exposure (when compared to 0h post-exposure; Fig 3B). This suggests that the effect of 5-aza-dC
186 persists at least 48h after exposure has stopped. Finally, 3,833 genes displayed significantly altered

187 methylation patterns in association with tissue type (Fig 3C), with lower methylation overall in the
188 head. However, we re-iterate that these main effects need to be contextualised within the pattern of
189 widespread interactions seen in the data, with 91% of genes exhibiting a significant second-order
190 interaction between exposure, tissue type and collection time.

191 Gene-set enrichment analysis, using *N. vitripennis* GO terms, when genes were ranked by their
192 response to exposure regime is presented in Table S2. Broadly, we identified 169 significantly enriched
193 GO terms ($p < 0.05$), and 80 of these terms fell under the "Biological process" (BP) category, 18 under
194 "Cellular component" (CC), and 71 under "Molecular function" (MF). The top five enriched GO terms
195 were "sequence-specific DNA binding" (GO:0043565), "regulation of transcription, DNA templated"
196 (GO:0006355), "ATP binding" (GO:0005524), "DNA binding transcription factor activity" (GO:
197 0003700), and "protein kinase activity" (GO:0004672) (Table S2). When the analysis was repeated
198 using GO terms for the *D. melanogaster* orthologs, a higher number of significantly enriched GO terms
199 was obtained (700; 538 BP, CC 71, MF 91; Table S3). The top five GO terms for the *D. melanogaster*
200 analysis were "nucleus" (GO:0005634), "positive regulation of transcription by RNA polymerase II"
201 (GO:0045944), "regulation of Notch signalling pathway" (GO:0008593), "imaginal disc-derived wing
202 morphogenesis" (GO:0007476), and, making a reappearance, "sequence-specific DNA binding"
203 (GO:0043565). This analysis combined GO-terms when *N. vitripennis* genes had multiple *D.*
204 *melanogaster* orthologs, which could lead to bias. H, however, we retrieve very similar enriched terms
205 when we used only the GO terms that were shared between the multiple orthologs (Table S4),
206 suggesting any bias caused by this approach is small.

207

208 **Discussion**

209 Our genome-wide analysis of CpGs across more than 11,000 genes showed that the demethylating
210 agent 5-aza-dC has very widespread effects on patterns of DNA methylation in the *Nasonia vitripennis*
211 genome. The methylome has been characterised in *Nasonia vitripennis* by Wang *et al.* (2013) and
212 Beeler *et al.* (2014), and in common with other insect species, DNA methylation appears to be
213 primarily associated with gene bodies (including intron-exon boundaries), and a rather low proportion
214 of CpGs are methylated. As mentioned above, previous work that has attempted to confirm the effects
215 of 5-aza-dC on DNA methylation have had mixed results. Pegoraro *et al.* (2016) confirmed the effects
216 of 5-aza-dC across five genes (four via a qPCR MethylQuant assay and one via cloning a fragment of
217 the *msn* gene from treated versus untreated individuals), whilst Ellers *et al.* (2019) have recently failed
218 to show any effects of 5-aza-dC on the methylation status of 155 CpGs across nine genes (they used
219 high-throughput bisulfite amplicon sequencing (BSAS); see Redshaw *et al.* (2014) for a discussion of

220 different methylation quantification assays). Our larger-scale study has both confirmed the action of
221 5-aza-dC as a disruptor of methylation in *Nasonia*, but also suggested why it might be so difficult to
222 pin down those effects. In particular, we find that whilst we see statistically significant main effects of
223 exposure regime on DNA methylation (which includes differences in length of exposure), we also see
224 effects of tissue type (albeit at the crude level of heads versus bodies), and time since exposure to 5-
225 aza-dC. Moreover, we also see many statistically significant interactions between these factors; all
226 told, the action of 5-aza-dC is likely to be context-dependent, both in terms of the sequences more
227 likely to be targeted but also in terms of the experimental details, including target tissues.
228 Nonetheless, our results strongly suggest that 5-aza-dC is not chemically inert in terms of influencing
229 DNA methylation in *Nasonia* – indeed, quite the opposite.

230 That the effects of 5-aza-dC are non-random in terms of the genes it influences are further
231 confirmed by our gene ontology enrichment analyses. We see enrichment for genes associated with
232 DNA binding, regulation of transcription, and transcription factor activity. These processes are of
233 interest given that we might expect the regulation of transcription to be a crucial component of the
234 epigenetic role of DNA methylation, especially given what we know from both mammals and insects,
235 even though the targets (throughout the genome versus exons) and outcomes (transcriptional
236 silencing versus stable gene expression) are typically different in the two taxa. Our analysis certainly
237 confirms the sequence-specific action of 5-aza-dC seen in other species.

238 A further important effect we see is that 5-aza-dC exposure can result in hyper-methylation
239 as well as hypo-methylation. This means that the characterisation of 5-aza-dC as a strictly
240 demethylating agent may be somewhat misleading. The complexity of outcomes of 5-aza-dC in
241 mammals, in particular in terms of trying to understand its action for cancer treatments, is becoming
242 increasingly well-appreciated (Seelan *et al.* 2018). As such, our data confirm in an insect what is being
243 seen in mammalian cells. We are not well placed to speculate on the mechanism(s) by which 5-aza-dC
244 may lead to hyper-methylation (see instead Seelan *et al.* 2018), even though its hypomethylating
245 effect is well characterised chemically (see Introduction). However, if 5-aza-dC disrupts the regulation
246 of the DNA methylation machinery itself, for instance in terms of how the Dnmt genes are expressed
247 and act, then the downstream outcomes could be both hyper- and hypo-methylation. Perhaps related
248 to this, in their study of DNA methylation and the control of the diapause response in *Nasonia*,
249 Pegoraro *et al.* (2016) used both 5-aza-dC and RNA interference (RNAi) to manipulate methylation, in
250 the latter case disrupting DNA methylation by knocking down both Dnmt1a and Dnmt3. Whilst both
251 exposure to 5-aza-dC and the RNAi constructs affected how diapause was influenced by exposure to
252 short or long day-length, they did so in different ways. Whilst both manipulations abolished the day-
253 length response, the RNAi of Dnmt1a and Dnmt3 resulted in a change in how wasps responded to

254 long-day length (increasing the diapause response), whilst 5-aza-dC led to a change in response to
255 short-day length (decreasing the diapause response). As such, different manipulations of DNA
256 methylation may lead to different phenotypic outcomes, depending on the regulatory causes and
257 consequences of the methylation pattern itself. More generally then, it is clear that much remains to
258 be elucidated in how 5-aza-dC shapes methylation patterns, in both insects and mammals.

259 On a more technical note, perhaps unsurprisingly we found more significant genes in the GO
260 enrichment analysis using the GO terms for *Drosophila melanogaster* orthologs. This is perhaps
261 because of the more detailed gene ontology information available for the *D. melanogaster* genome.
262 Nonetheless, the two analyses provide a similar story in terms of GO enrichment. In addition, the
263 majority of our analyses were undertaken at the level of the gene. Perhaps given the variation in
264 coverage at the level of individual CpGs, the clustering analyses were not by any means as clear cut as
265 the gene-level results. However, in the t-SNE analysis (Fig S1), the control replicates cluster together,
266 albeit within the broader distribution of the other two exposure treatment groups. The effect of
267 genomic and sequence context may be particularly relevant here: our analysis at the level of the gene
268 (see Methods for details) brings together CpGs that share a lot of context, by virtue of being associated
269 with a given genetic locus. The differences in the clustering at the individual CpG versus gene levels
270 may therefore represent the fact that the action of 5-aza-dC does appear to be non-random, and that
271 we gain clarity when we consider closely associated CpGs whose patterns of methylation, and
272 responses to 5-aza-dC, are more similar than randomly chosen CpGs across the genome.

273 In conclusion, our results suggest that 5-aza-dC remains a potential tool for evolutionary
274 ecologists wishing to explore how DNA methylation influences phenotypic expression. As with all
275 genomic approaches however, no one approach gives all the answers. 5-aza-dC may prove a very
276 useful, if rather blunt, tool for first assessing the range of phenotypes that are influenced by
277 methylation status, whilst no doubt more targeted knock-downs and genetic manipulations will be
278 needed to link phenotypes to causal CpGs.

279

280 **Materials and Methods**

281 *Study species*

282 *Nasonia vitripennis* (Hymenoptera, Chalcicoidea) is a generalist parasitoid of large dipteran pupae
283 including species of Calliphoridae. Females oviposit between 20 and 50 eggs in an individual host, with
284 male offspring emerging just before females (after approximately 14 d at 25 °C: Whiting 1967). Males
285 are brachypterous and unable to fly, remaining close to the emergence site where they compete with

286 each other for emerging females, including their sisters. Females disperse after mating to locate new
287 hosts. The females used in this experiment were from the wild-type AsymC strain, originally isolated
288 in 1986 by curing the wild-type strain LabII of *Wolbachia*. Wasps have been maintained on *Calliphora*
289 *vomitoria* or *C. vicina* hosts at 25 °C, 16L:8D light conditions ever since, and AsymC is also the reference
290 genome strain (Werren *et al*. 2010). The quantitative genetic basis of sex ratio variation has been
291 repeatedly quantified (Pannebakker *et al.* 2008, 2011), whilst whole-genome transcription studies
292 have confirmed that facultative sex allocation takes place without discernible changes in gene
293 expression (Cook *et al.* 2015b, 2018).

294

295 *5-aza-2'-deoxycytidine*

296 We used 5-aza-dC (Sigma Aldrich Company Ltd., Gillingham, UK) at a non-lethal concentration (10 µM).
297 Widely used in epigenetic anticancer treatments (Seelan *et al.* 2018), 5-aza-dC is a nucleoside
298 analogue that works by inhibiting DNA methyltransferase activity (Piskala and Sorm 1964, Jones and
299 Taylor 1980, Momparler 1985, Christman 2002). When initially introduced to a cell, 5-aza-dC is
300 essentially inert but is then converted by processes inherent in the cell to its active form 5-aza-dCTP
301 (Seelan *et al.* 2018). This active form is readily incorporated into DNA during S-phase of the cell cycle
302 in the place of cytosine. In the daughter strands arising from a replication event, DNA
303 methyltransferase 1 (Dnmt1) will attempt to restore the patterns of methylation that were present in
304 the parent strand. However, Dnmt1 becomes irreversibly bound to the 5-aza-dCTP residues in the
305 daughter strands, resulting in a diminishing pool of Dnmt1. Thus, administering 5-aza-dC results in a
306 passive loss of DNA methylation, the key effect of 5-aza-dC (for a fuller description, see Cook *et al.*
307 2019).

308

309 *Experimental design*

310 To control for possible host and other maternal effects, experimental females were not drawn straight
311 from stock populations. Instead, two-day-old, mated, wild-type AsymC females were isolated from
312 the mass cultures into individual glass vials. Each female was provided with three hosts and allowed
313 to oviposit. Experimental females were drawn from the F1 generation, two days after emergence, one
314 female per "grandmother". We then employed a three by three factorial design whereby females
315 were allocated to one of three 5-aza-dC exposure regimes and were harvested post-feeding at three
316 time points. Exposure regimes were a) 20% sucrose for 24h, b) 20% sucrose supplemented with 10
317 µM 5-aza-dC for 24h and c) 20% sucrose supplemented with 10 µM 5-aza-dC for 48h. For each of the

318 three regimes, females were harvested at 0, 24 or 48h post-exposure. Throughout the experiment,
319 females were kept individually in glass vials and incubated at 25 °C, 16L :8D. Solutions (200 µl) were
320 provided in the lid of a 2 ml microcentrifuge tube. For individuals exposed to 5-aza-dC for 48 h, the
321 solutions were topped up by 100µl half-way through the exposure period to compensate for
322 evaporation. All solutions were stained green with food colouring such that a subset of 10 females
323 from each treatment combination could be dissected to check that feeding had occurred. At the point
324 of harvest, females were flash-frozen in liquid nitrogen and stored at -80 °C prior to DNA extraction.

325

326 *DNA extraction and bisulphite-conversion*

327 Prior to DNA extraction, individuals were removed from -80 °C storage and heads were quickly excised
328 from bodies using a sterile scalpel and a cryloyser to keep samples cold. Heads (and likewise bodies)
329 from the same treatment combination were randomly pooled into groups of 10 to give a maximum of
330 3 biological replicates for each treatment combination giving a total of 54 DNA samples. Briefly, tissue
331 was homogenised using a micropestle in 350 µl CTAB (Hexadecyltrimethylammonium Bromide) buffer
332 with subsequent overnight proteinase K digestion (400 µg) at 56 °C. Samples were cooled to room
333 temperature prior to RNase A digestion for 1 hour at 37 °C. A chloroform: isoamyl-alcohol wash (300
334 µl) was performed twice before ethanol/sodium acetate precipitation of the DNA. DNA purity was
335 analysed using the Nanodrop spectrophotometer with all samples having a 260/280 ratio of ≥1.8. DNA
336 integrity was confirmed using agarose gel electrophoresis.

337

338 *Bisulphite sequencing*

339 Bisulphite conversion, library preparation and sequencing were carried out by Edinburgh Genomics.
340 Briefly, bisulphite conversion was carried out using the EZ DNA Methylation Kit (Zymo Research, Irvine
341 USA), designed to reduce template degradation during the harsh bisulphite treatment whilst ensuring
342 complete conversion of the DNA. Library preparation was carried out using the TruSeq DNA
343 methylation Kit which takes the single-stranded DNA that results from bisulphite treatment and
344 converts it into an Illumina sequencing library. Sequencing of bisulphite-treated DNA was carried out
345 on the Illumina HiSeq 2500 in high output mode with 125bp paired-end reads. Library preparation
346 failed for three of the samples due to low levels of input DNA, leaving only two biological replicates
347 for three out of 18 treatment combinations and a total of 51 Bisulphite-sequenced samples (Table 1).

348

349 *Raw data processing*

350 Initial data processing was carried out by Edinburgh Genomics. Briefly, reads were trimmed using
351 Cutadapt version 1.12 removing any adapter sequences and for quality using a threshold of 30. After
352 trimming reads were required to have a minimum length of 50 bp. *Nasonia vitripennis* genome version
353 2.1 and the associated annotation, available for download at Ensembl
354 (http://metazoa.ensembl.org/Nasonia_vitripennis/Info/Index), were used as a reference. Reads were
355 aligned to the reference genome using Bismark (version 0.16) with the parameter "bowtie2". PCR
356 duplicates were removed from the resulting BAM files using samtools with the parameter "view -F
357 1024". Data for each CpG was then extracted from the BAM files using the Bioconductor package
358 methylKit (version 0.99.3) with R (version 3.3.1; R Core Team 2017). The methylation status of each
359 CpG for all samples was calculated using the function "processBismarkAln" using the default
360 parameters, except "min.cov=1 no;ap=TRUE", which excludes CpG's not covered by any reads across
361 all samples and does not count CpGs covered by overlapping ends of a paired read as covered twice.
362 The 1% of bases with the highest coverage were removed using the function "filterByCoverage" to
363 eliminate potential PCR bias. After initial processing as outlined above, the raw data comprised
364 percentage methylation at 4,150,376 CpG loci (no. of reads in which the CpG is methylated/total no.
365 of reads covering the CpG) for each of 51 samples. CpG read counts were then summed by gene,
366 (including 1Kb up and downstream) using gene annotation from Ensembl (version 2.1. available from:
367 ftp://ftp.ensemblgenomes.org/pub/metazoa/release-40/gff3/nasonia_vitripennis). Around half
368 (2,119,643) of the CpG's in the full dataset fell within genic regions with 14,765/17,279 genes (85%)
369 having at least one CpG present. Median coverage across the whole experiment (N = 51 BS-seq
370 libraries) was 898X, but coverage was heterogeneous between loci (Fig 4) and samples (Table S5).
371 Mean read coverage per sample ranged from 2.6X to 48.2X. Four samples had less than 10X mean
372 coverage per CpG, however, these samples were distributed across treatment groups.

373

374 *Statistical analysis*

375 The overall structure of our dataset in terms of the proportion of methylated reads per gene and per
376 loci was visualised with two complementary analyses: a principle component analysis (PCA), which
377 allows samples to be visualised in relation to the two major axes of variation in our experiment, and a
378 T-distributed Stochastic Neighbour Embedding (t-SNE) analysis. t-SNE is a machine-learning algorithm
379 for data visualisation well-suited for high-dimensional datasets (van der Maaten and Hinton 2008).
380 The PCA was conducted using the R base function prcomp with default values. To perform the t-SNE
381 analysis we used the R package Rtsne (Krijthe 2015), with default values and perplexity set to 15. For

382 the PCA and t-SNE analyses any genes or loci with 0 coverage in any one sample were discarded prior
383 to analysis.

384 To examine the effect of 5-aza-dC exposure on individual genes, we modelled the number of
385 methylated and unmethylated reads from a gene as a generalized linear model (GLM) with a binomial
386 distribution with the following terms: exposure, tissue-type, collection time and their interactions.
387 Genes with 0 coverage for all samples in any particular treatment group were excluded from our
388 analysis (leaving a total of 11582 genes). Significance of each term was determined by a likelihood
389 ratio test, which was corrected for multiple testing using Benjamini and Hochberg's algorithm
390 (Benjamini and Hochberg 1995), with statistical significance set to 5%. Coefficients (log odds of
391 methylation) for each term were extracted from the full model. All statistical analyses were conducted
392 in R (v. 3.4.1; R Core Team 2017).

393

394 *Functional enrichment*

395 *N. vitripennis* GO-terms were downloaded from "hymenopteramine"
396 (http://hymenopteragenome.org/hymenopteramine/begin.do) on 27th September 2018. In addition,
397 we used hymenopteramine to firstly obtain *D. melanogaster* orthologs for each *N. vitripennis* gene,
398 from which we could then obtain the associated *D. melanogaster* GO terms. We used both of these
399 GO-term sets for enrichment analyses (below). Some *N. vitripennis* genes had multiple *D.*
400 *melanogaster* orthologs. In these cases, all GO-terms from each *D. melanogaster* ortholog were
401 combined together. To ensure this did not bias our results we repeated our analyses, but instead of
402 combining GO-terms we kept only GO-terms that were shared between all orthologs.

403 To examine which processes are most affected by 5-aza-dC exposure, we conducted gene set
404 enrichment analyses using the R package TopGO (v. 2.28.0; Rahnenfuhrer 2016) using the elim
405 algorithm to account for the GO topology. Gene set enrichment analyses identify enriched GO terms
406 in a threshold-free way, by finding GO-terms that are overrepresented at the top of a ranked list of
407 genes. Here we ranked genes by FDR for the exposure effect. GO terms were considered to be
408 significantly enriched when $p < 0.05$.

409

410 **Acknowledgments**

411 This work was initiated as part of NERC grant NE/J024481/1. We are very grateful to Mark Lammers
412 and Jacintha Ellers for their collegiate and constructive engagement with our previous DNA

methylation work. We are also grateful to Matthew Arno and the Edinburgh Genomics team for their sequencing and downstream support. Raw data are available from the Gene Expression Omnibus hosted by the National Center for Biotechnology Information (Accession: GSE125388). Scripts for the analyses in this paper are available at: https://github.com/DarrenJParker/5-Aza-dC_methylation.

444 **Tables**

445 **Table 1. A summary of the experimental design[a].**

Treatment Combination	5-aza-dC exposure regime	Harvest time	Tissue type	No. replicates (1 x rep. = 1 pool of 10)
1	20% sucrose for 24 h	0h	Bodies	3
2	20% sucrose for 24 h	0h	Heads	3
3	20% sucrose for 24 h	24h	Bodies	3
4	20% sucrose for 24 h	24h	Heads	3
5	20% sucrose for 24 h	48h	Bodies	3
6	20% sucrose for 24 h	48h	Heads	3
7	10µM 5-aza-dC for 24 h	0h	Bodies	3
8	10µM 5-aza-dC for 24 h	0h	Heads	3
9	10µM 5-aza-dC for 24 h	24h	Bodies	3
10	10µM 5-aza-dC for 24 h	24h	Heads	2
11	10µM 5-aza-dC for 24 h	48h	Bodies	3
12	10µM 5-aza-dC for 24 h	48h	Heads	2
13	10µM 5-aza-dC for 48 h	0h	Bodies	3
14	10µM 5-aza-dC for 48 h	0h	Heads	2
15	10µM 5-aza-dC for 48 h	24h	Bodies	3
16	10µM 5-aza-dC for 48 h	24h	Heads	3
17	10µM 5-aza-dC for 48 h	48h	Bodies	3
18	10µM 5-aza-dC for 48 h	48h	Heads	3

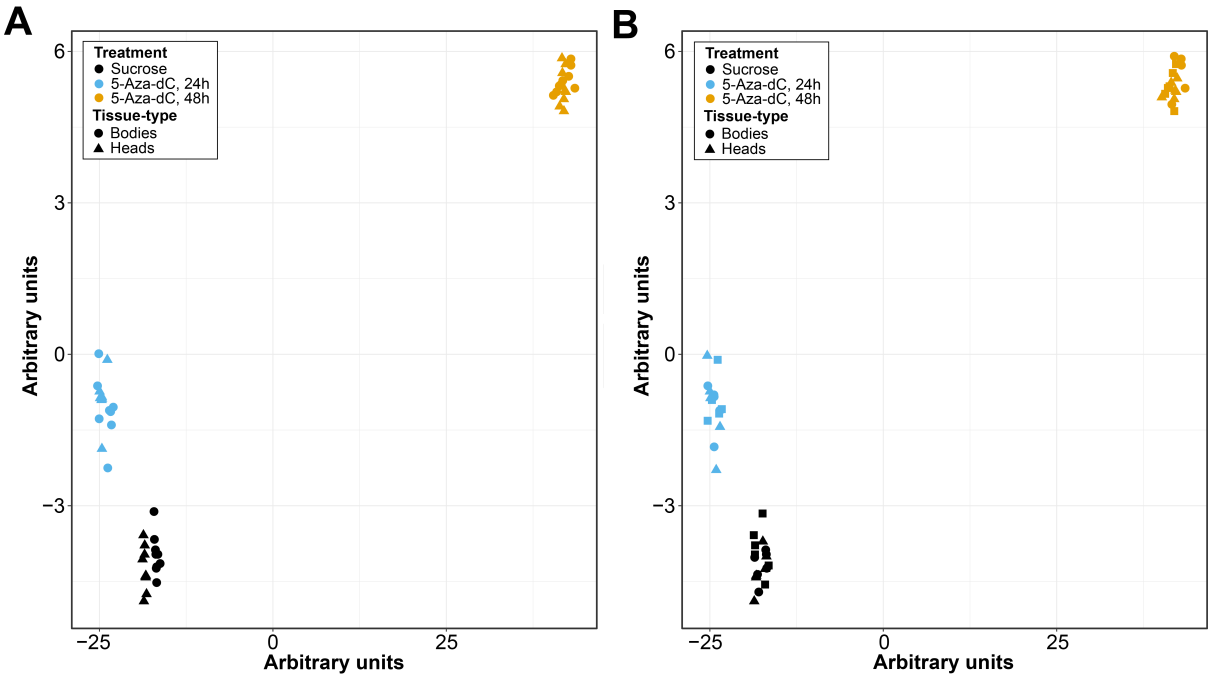
446 [a] Female wasps were exposed to 1 of 3 5-aza-dC exposure regimes including a control. Females were
447 then harvested at 1 of 3 timepoints after the exposure period heads were excised from bodies to
448 give a degree of tissue-specificity to the analysis. Three biological replicates were available for
449 almost all treatment combinations with each biological replicate consisting of tissue from three *N.*
450 *vitripennis* females.

451

452

453 **Figures**

454



455

456 **Fig 1. t-SNE of the proportion of methylated reads per gene.** Shows the effects of (A) Exposure

457 regime and tissue type and (B) Exposure regime and harvest time.

458

459

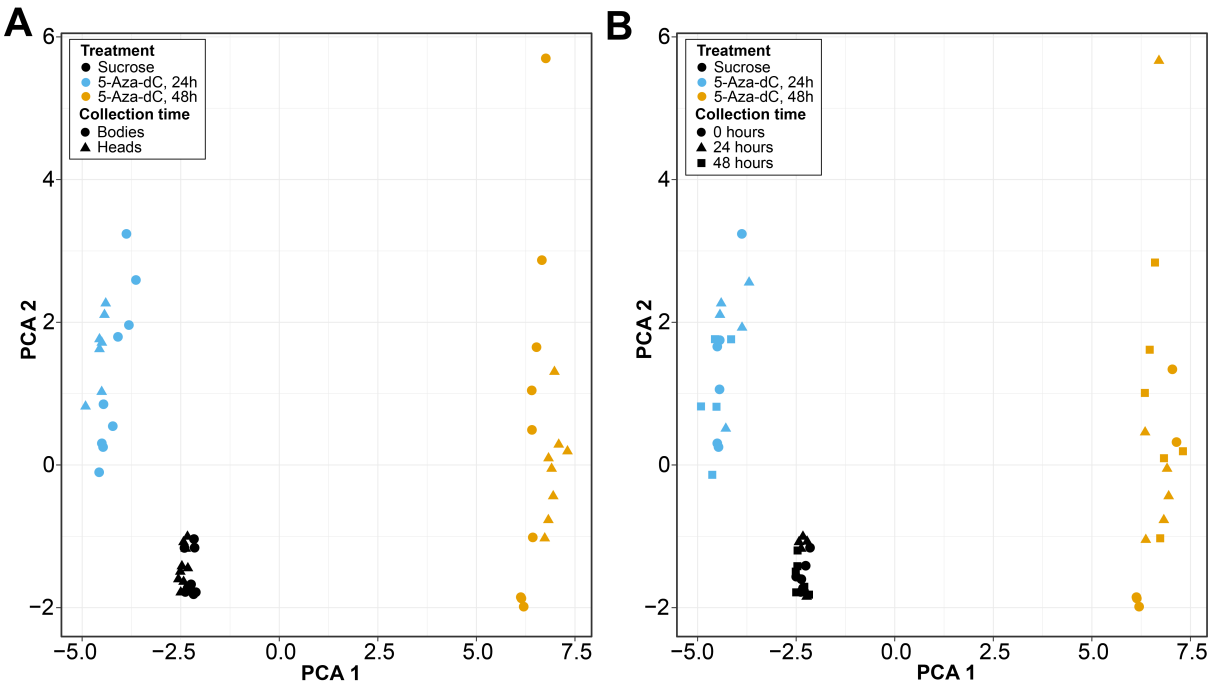


Fig 2. PCA of the proportion of methylated reads per gene. Shows the effects of (A) Exposure regime and tissue type and (B) Exposure regime and harvest time.

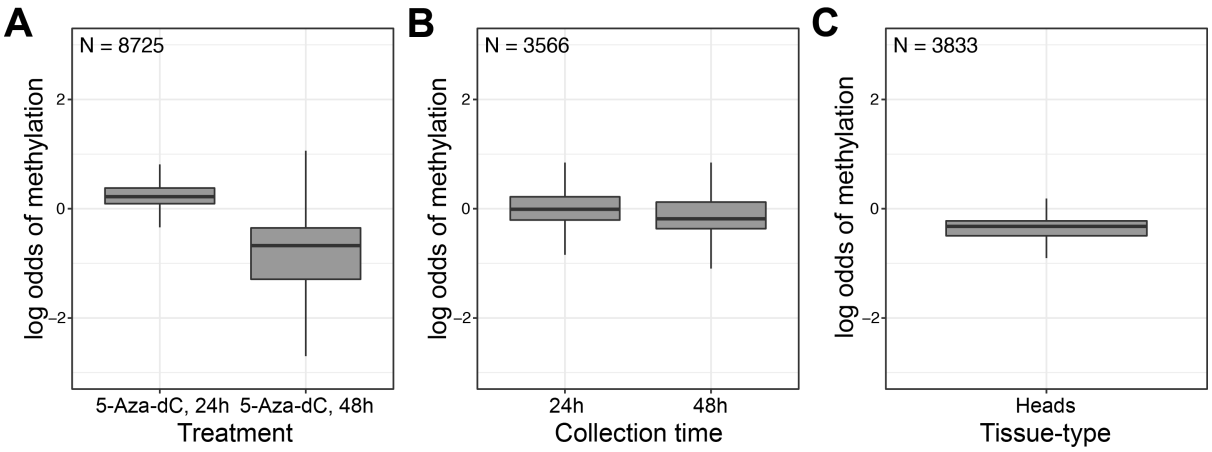


Fig 3. Gene methylation status displayed as log odds of methylation for each main effect in the experiment. A significant effect on gene methylation status was noted in response to (A) 5-aza-dC exposure regime (relative to sucrose control), (B) harvest time after exposure (relative to zero hours) and (C) tissue-type (heads relative to bodies). N= number of genes with significantly differentially methylated CpGs in association with each main effect.

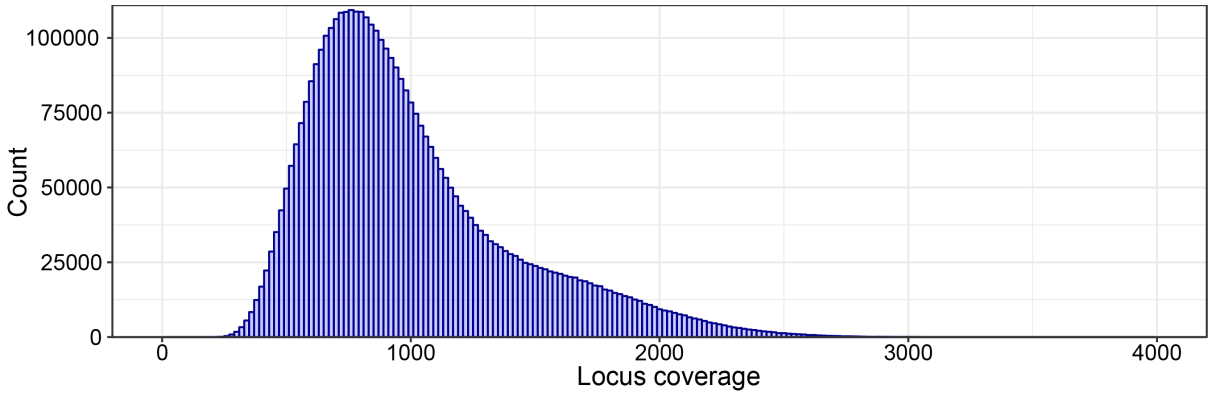


Fig 4. Read coverage by locus across all 51 BS-Seq libraries.

References

Amarasinghe HE, Clayton CI, Mallon EB. 2014. Methylation and worker reproduction in the bumble-bee (*Bombus terrestris*). *Proc R Soc Lond B Biol Sci.* 281: 20132502.

Ashbrook DG, Mulligan MK, Williams RW. 2018. Post-genomic behavioural genetics: From revolution to routine. *Genes Brain Behav.* 17: e12441.

Beeler SM, Wong GT, Zheng JM, Bush EC, Remnant EJ, Oldroyd DP, Drewell RA. 2014. Whole-genome DNA methylation profile of the Jewel Wasp (Nasonia vitripennis). G3 (Bethesda) 4: 383-388.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol.* 57: 289–300.

Bewick AJ, Vogel KJ, Moore AJ, Schmitz RJ. 2017. Evolution of DNA methylation across Insects. *Mol Biol Evol.* 34: 654-665.

Biémont C. 2010. From genotype to phenotype. What do epigenetics and epigenomics tell us? *Heredity* 105: 1-3.

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

Bonduriansky R, Day T. 2018. Extended Heredity. Princeton: Princeton University Press.

Bryan J, Kantarjian H, Garcia-Manero G, Jabbour E. 2011. Pharmacokinetic evaluation of decitabine for the treatment of leukemia. *Expert Opin Drug Metab Toxicol.* 7: 661-672.

Burggren W. 2016. Epigenetic Inheritance and its role in evolutionary biology: re-evaluation and new perspectives. *Biology.* 5: 24.

Burton-Chellew MN, Koevoets T, Grillenberger BK, Sykes EM, Underwood SL, Bijlsma K, Gadau J, van de Zande L, Beukeboom LW, West SA, *et al.* 2008. Facultative sex ratio adjustment in natural populations of wasps: cues of local mate competition. *Am Nat.* 172: 393-404.

Christman JK. 2002. 5-azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. *Oncogene* 21: 5483-5495.

Constancia M, Hemberger M, Hughes J, Dean W, Ferguson-Smith A, Fundele R, Stewart F, Kelsey G, Fowden A, Sibley C, *et al.* 2002. Placental-specific IGF-II is a major modulator of placental and fetal growth. *Nature* 417(6892): 945-948.

Cook N, Pannebakker BA, Tauber E, Shuker DM. 2015a. DNA methylation and sex allocation in the parasitoid wasp Nasonia vitripennis. *Am Nat.* 186: 513-518.

Cook N, Trivedi U, Pannebakker BA, Blaxter M, Ritchie MG, Tauber E, Sneddon T, Shuker DM. 2015b. Oviposition but not sex allocation is associated with transcriptomic changes in females of the parasitoid wasp *Nasonia vitripennis*. *G3 (Bethesda)* 5: 2885-2892.

Cook N, Boulton RA, Green J, Trivedi U, Tauber E, Pannebakker BA, Ritchie MG, Shuker DM. 2018. Differential gene expression is not required for facultative sex allocation: a transcriptome analysis of brain tissue in the parasitoid wasp *Nasonia vitripennis*. *Roy Soc Open Sci*. 5: 171718.

527
528 Cook N, Parker DJ, Tauber E, Pannebakker BA, Shuker DM. 2019. Validating the demethylating effects
529 of 5-aza-2'-deoxycytidine in insects requires a whole-genome approach (A reply to Ellers *et al.*). *Am*
530 *Nat.* Forthcoming.
531
532 Ellers J, Visser M, Mariën J, Kraaijeveld K, Lammers M. 2019. The importance of validating the
533 demethylating effect of 5-aza-dC in model species. *Am Nat*. Forthcoming.
534
535 Feng S, Cokus SJ, Zhang X, Chen P, Bostick M, Goll MG, Hetzel J, Jain J, Strauss SH, Halpern ME, *et al.*
536 2010. Conservation and divergence of methylation patterning in plants and animals. *Proc Natl Acad*
537 *Sci USA.* 107: 8689-8694.
538
539 Glastad KM, Hunt BG, Yi SV, Goodisman MAD. 2011. DNA methylation in insects: on the brink of the
540 epigenomic era. *Insect Mol Biol.* 20: 553-65.
541
542 Hagemann S, Heil O, Lyko F, Brueckner B. 2011. Azacytidine and decitabine induce gene-specific and
543 non-random DNA demethylation in human cancer cell lines. *PLoS One.* 6: e17388.
544
545 Haig D. 1997. Parental antagonism, relatedness asymmetries and genomic imprinting. *Proc R Soc Lond*
546 *B Biol Sci.* 264: 1657-1662.
547
548 Haig D. 2002. Genomic imprinting and kinship. New Brunswick: Rutgers University Press.
549
550 Hamilton WD. 1967. Extraordinary sex ratios. *Science* 156: 477-488.
551
552 Jones PA, Taylor SM. 1980. Cellular differentiation, cytidine analogs and DNA methylation. *Cell* 20: 85-
553 93.
554
555 Krijthe JH. 2015. Rtsne: T-Distributed Stochastic Neighbor Embedding using a Barnes-Hut
556 Implementation. Available from: https://github.com/jkrijthe/Rtsne
557
558 Kucharski R, Maleszka J, Foret S, Maleszka R. 2008. Nutritional control of reproductive status in
559 honeybees via DNA methylation. *Science* 319: 1827-1830.
560
561 Lind MI, Spagopoulou F. 2018. Evolutionary consequences of epigenetic inheritance. *Heredity.* 121:
562 205-209.
563
564 Lyko F, Foret S, Kucharski R, Wolf S, Falckenhayn C, Maleszka R. 2010. The honey bee epigenomes:
565 differential methylation of brain DNA in queens and workers. *PLoS Biol.* 8: e1000506.
566
567 Lyko F. The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. 2018. *Nat Rev*
568 *Genet.* 19: 81-92.
569
570 Momparler RL. Molecular, cellular and animal pharmacology of 5-aza-2'-deoxycytidine. 1985.
571 *Pharmacol Ther.* 30: 287-299.
572
573 Momparler RL. 2012. A perspective on the comparative antileukemic activity of 5-aza-2'-deoxycytidine
574 (Decitabine) and 5-azacytidine (Vidaza). *Pharmaceuticals* (Basel). 5: 875-881.
575
576 Moore T, Haig D. 1991. Genomic imprinting in mammalian development: a parental tug-of-war. Trends
577 Genet. 7: 45-49.

578
579 Mossman D, Kim K, Scott RJ. 2010. Demethylation by 5-aza-2'-deoxycytidine in colorectal cancer cells
580 targets genomic DNA whilst promoter CpG island methylation persists. *BMC Cancer* 10: 366.
581
582 Pannebakker BA, Halligan DL, Reynolds KT, Ballantyne GA, Shuker DM, West SA. 2008. Effects of
583 spontaneous mutation accumulation on sex ratio traits. *Evolution*. 62: 1921-1935.
584
585 Pannebakker BA, Watt R, Knott SA, West SA, Shuker DM. 2011. The quantitative genetic basis of sex
586 ratio variation in *Nasonia vitripennis*: a QTL study. *J Evol Biol.* 24: 12-22.
587
588 Patten MM, Ross L, Curley JP, Queller DC, Bonduriansky R. Wolf JB. 2014. The evolution of genomic
589 imprinting: theories, predictions and empirical tests. *Heredity* 113: 119-128.
590
591 Pegoraro M, Bafna A, Davies NJ, Shuker DM, Tauber E. 2016. DNA methylation changes induced by
592 long and short photoperiods in Nasonia. *Genome Res.* 26: 203-210.
593
594 Piskala, Sorm F. 1964. Nucleic acid components and their analogues. LI. Synthesis of 1-glycosyl
595 derivatives of 5-azauracil and 5-azacytosine. *Collect Czechoslov Chem Commun.* 29: 2060-2076.
596
597 R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for
598 Statistical Computing, Vienna, Austria. Available from https://www.R-project.org
599
600 Rahnenfuhrer AAJ. 2016. *topGO: Enrichment Analysis for Gene Ontology*. R package version 2.28.0.
601 Available from http://bioconductor.org/packages/release/bioc/html/topGO.html
602
603 Ramos M, Wijetunga NA, McLellan AS, Suzuki M, Greally JM. 2015. DNA demethylation by 5-aza-2'-
604 deoxycytidine is imprinted, targeted to euchromatin and has limited transcriptional consequences.
605 *Epigenetics Chromatin.* 8:11.
606
607 Redshaw N, Huggett JF, Taylor MS, Foy CA, Devonshire AS. 2014. Quantification of epigenetic
608 biomarkers: an evaluation of established and emerging methods for DNA methylation analysis. *BMC*
609 *Genomics* 15: 1174.
610
611 Schield DR, Walsh MR, Card DC, Andrew AL, Adams RH, Castoe TA. 2016. EpiRADseq: scalable analysis
612 of genomewide patters of methylation using next-generation sequencing. *Meth Ecol Evol.* 7: 60-69.
613
614 Seelan RS, Mukhopadhyay P, Pisano MM, Greene RM. 2018. Effects of 5-Aza-2'-deoxycytidine
615 (decitabine) on gene expression. *Drug Metab Rev.* 50: 193-207.
616
617 Shuker DM, West SA. 2004. Information constraints and the precision of adaptation: sex ratio
618 manipulation in wasps. *Proc Natl Acad Sci USA.* 101: 10363-10367.
619
620 Smith G, Smith C, Kenny JG, Chaudhuri RR, Ritchie MG. 2015. Genome-wide DNA methylation patterns
621 in wild samples of two morphotypes of Threespine Stickleback (Gasterosteus aculeatus). *Mol Biol Evol.*
622 32: 888-895.
623
624 Taylor PD, Bulmer MG. 1980. Local Mate Competition and the sex-ratio. *J Theoret Biol.* 86: 409-419.
625 van der Maaten LJP, Hinton GE. 2008. Visualizing high-dimensional data using t-SNE. *J Mach Learn Res.*
626 9: 2579-2605.
627

628 Wang X, Wheeler D, Avery A, Rago A, Choi JH. 2013. Function and evolution of DNA methylation in
629 *Nasonia vitripennis*. *PLoS Genet*. 9: e1003872.
630
631 Wang X, Werren JH, Clark AG. 2016. Allele-specific transcriptome and methylome analysis reveals
632 stable inheritance and cis-regulation of DNA methylation in Nasonia. *PLoS Biol.* 14: e1002500.
633
634 Werren JH. 1980. Sex ratio adaptations to local mate competition in a parasitic wasp. *Science* 208:
635 1157-1159.
636
637 Werren JH. 1983. Sex ratio evolution under local mate competition in a parasitic wasp. *Evolution* 37:
638 116-124.
639
640 Werren JH, *et al.* 2010. Functional and evolutionary insights from the genomes of three parasitoid
641 *Nasonia* species. *Science* 327: 343-348.
642
643 West SA. 2009. Sex allocation. Princeton: Princeton University Press.
644
645 Whiting AR. 1967. The biology of the parasitic wasp *Mormoniella vitripennis* [=*Nasonia brevicornis*]
646 (Walker). *Q Rev Biol.* 42: 333-406.
647
648 Wild G, West SA. 2008. Genomic imprinting and sex allocation. *Am Nat.* 173: E1-14.
649
650 Zemach A, McDaniel IE, Silva P, Zilberman D. 2010. Genome-wide evolutionary analysis of eukaryotic
651 DNA methylation. *Science* 328: 916-919.
652