



**HAL**  
open science

## Frequency and mitotic heritability of epimutations in *Schistosoma mansoni*

David Roquis, Anne Rognon, Cristian Chaparro, Jérôme Boissier, Nathalie Arancibia, Celine Cosseau, Hugues Parrinello, Christoph Grunau

► **To cite this version:**

David Roquis, Anne Rognon, Cristian Chaparro, Jérôme Boissier, Nathalie Arancibia, et al.. Frequency and mitotic heritability of epimutations in *Schistosoma mansoni*. *Molecular Ecology*, 2016, EPIGENETIC STUDIES IN ECOLOGY AND EVOLUTION, 25 (8), pp.1741-1758. 10.1111/mec.13555 . hal-01311089

**HAL Id: hal-01311089**

**<https://sde.hal.science/hal-01311089>**

Submitted on 1 Apr 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# 1 Frequency and mitotic heritability of epimutations in *Schistosoma* 2 *mansoni*

3  
4 David Roquis<sup>1,2</sup>, Anne Rognon<sup>1,2</sup>, Cristian Chapparo, Jerome Boissier<sup>1,2</sup>, Nathalie Arancibia<sup>1,2</sup>, Celine  
5 Cosseau<sup>1,2</sup>, Hugues Parrinello<sup>3</sup> & Christoph Grunau<sup>1,2</sup>

6  
7 1 Université de Perpignan Via Domitia, Perpignan, F-66860, France,  
8 2 CNRS, UMR 5244, Interactions Hôtes-Pathogènes-Environnements (IHPE), Perpignan, F-66860, France,  
9 3 CNRS, GenomiX IBISA, Montpellier, France,  
10

## 11 Abstract

12  
13 *Schistosoma mansoni* is a parasitic platyhelminth responsible for intestinal bilharzia. It has a  
14 complex life cycle, infecting a freshwater snail of the *Biomphalaria* genus, and then a mammalian  
15 host. *S. mansoni* adapts rapidly to new (allopatric) strains of its intermediate host. To study the  
16 importance of epimutations in this process we infected sympatric and allopatric mollusk strains  
17 with parasite clones. CHIP-Seq was done on four histone modifications (H3K4me3, H3K27me3,  
18 H3K27ac and H4K20me1) in parallel with genome-wide DNA resequencing (i) on parasite larvae  
19 shed by the infected snails, and (ii) on adult worms that had developed from the larvae. No change  
20 in single nucleotide polymorphisms (SNP) and no mobilization of transposable elements (TE) was  
21 observed, but 58 – 105 copy number variations (CNV) within the parasite clones in different  
22 mollusks were detected. We also observed that the allopatric environment induces three types of  
23 chromatin structure changes: (i) host-induced changes on larvae epigenomes in 51 regions of the  
24 genome that are independent of the parasites' genetic background, (ii) spontaneous changes (not  
25 related to experimental condition or genotype of the parasite) at 64 locations, and (iii) 64  
26 chromatin structure differences dependent on the parasite genotype. Up to 45 % of the  
27 spontaneous, but none of the host-induced chromatin structure changes were transmitted to  
28 adults. In our model, the environment induces epigenetic changes at specific loci but only  
29 spontaneous epimutations are mitotically heritable and have therefore the potential to contribute  
30 to transgenerational inheritance. We also show that CNV are the only source of genetic variation  
31 and occur at the same order of magnitude as epimutations.  
32

## 33 Introduction

34

35 Adaptive evolution relies on the generation of heritable phenotypic variations. There is now an  
36 intense scientific discussion about the source of these variants (Laland *et al.* 2014). During the last  
37 couple of years, it has become clear that there are considerable non-genetic sources of heritable  
38 phenotypes (Jablonka *et al.* 1992; Jablonka & Lamb 2008; Danchin *et al.* 2011). In particular  
39 epigenetic modifications, that we define here tentatively as heritable changes in chromatin  
40 structure that can have an impact on gene expression, are in the center of a lively scientific debate  
41 (Kovalchuk 2011). Particularities of epimutations are (I) their metastability, *i.e.* changes in  
42 chromatin structure can be reversible and (II) their relations with the environment. In the genetic  
43 inheritance system, the environment acts solely as selective agent, but in the epigenetic system it  
44 could play a dual role, acting as an inducer of epimutations (Guerrero-Bosagna & Skinner 2011;  
45 Skinner *et al.* 2015) as well as applying selection on the resulting phenotypes. The question  
46 whether epimutations can generate new phenotypes is solved (Peaston & Whitelaw 2006; Nilsson  
47 & Skinner 2015). There are now many examples for phenotypes that are entirely or partially  
48 encoded by epialleles presenting changes in DNA methylation or histone modifications (Duncan *et al.*  
49 2014). The role of the environment, however, is far from being comprehended. There is  
50 empirical support that epimutations are generated spontaneously, independently of environment  
51 but at a higher frequency than genetic mutations (van der Graaf *et al.* 2015). It was also argued  
52 that environmental stress could provoke adaptive and heritable changes through targeted  
53 épimutations. Others concluded that stress leads to an increase of the epimutation rate, thus  
54 representing an ephemeral, but inducible version of the genetic inheritance system, *i.e.* a way to  
55 induce reversible but possibly inheritable phenotypic changes (Jiang *et al.* 2014). The objective of  
56 the present study was to determine the frequency and nature of histone-based epimutations  
57 during a change of environment, and to measure their mitotic heritability. To address this question  
58 we used a parasite-host system. In these systems reciprocal selective forces, such as infectivity for  
59 the parasite and resistance for the host, are strong and adaptive evolution is expected to be fast  
60 (Thrall *et al.* 2007; Tellier *et al.* 2014).

61

62 The plathelminth *Schistosoma mansoni* is the causing agent of intestinal schistosomiasis, a human  
63 disease affecting more than 67 million people worldwide (Oliveira *et al.* 2004; King 2010). Its  
64 complex life cycle (**Supplementary Figure 1**) starts with eggs hatching in fresh water. A free-  
65 swimming larvae (miracidium) is released and will enter the tegument of the intermediate mollusk  
66 host of the *Biomphalaria* genus. Miracidia quickly transform into primary sporocysts, multiply  
67 asexually and differentiate into secondary sporocysts. They generate multiple cercariae, another  
68 free-swimming larvae, which leave the mollusk, infect the definitive mammalian host (rodent or  
69 primate), and transform into schistosomula. They follow a complex maturation and migration  
70 process in the mammalian body to reach the adult worm stage. Male and female worms form  
71 couples in the mesenteric veins, where eggs are released, go through the intestine wall and are  
72 excreted in the feces. *S. mansoni* can infect a wide range of intermediate and definitive host  
73 species (Martins 1957; DeJong *et al.* 2001). Changes in the intermediate host strain and species  
74 has a reproducible phenotypic impact on the parasite, and more specifically on its life history traits  
75 (sex-ratio, growth, prevalence and intensity of the infection) (Dias *et al.* 1988; Lepesant *et al.*

76 2013; Theron *et al.* 2014), but the underlying mechanisms are still unknown. Compatibility  
77 polymorphism is classically used to explain infectivity between *S. mansoni* and *B. glabrata* (Roger  
78 *et al.* 2008; Mitta *et al.* 2012). However, recent work has uncovered that epigenetic modifications  
79 could also play a role in the adaptation to new host strains (Knight *et al.* 2013). Knowing this, we  
80 were interested to find out what would be the impact of different host strains on the genome and  
81 epigenome of *S. mansoni*. In order to disentangle the impact of epimutations and mutations, we  
82 decided to work with genetically identical parasites and to let them develop in different snails.  
83 Miracidia, coming from sexual reproduction, are not clonal, and although it is relatively easy to  
84 infect snails with sibling miracidia with low genetic diversity between them, there will still be  
85 genetic diversity. To overcome this problem, we used sporocyst grafts (Jourdane & Theron 1980;  
86 Jourdane 1984). This technique allows, after a mono-miracidium infection, to extract secondary  
87 sporocysts and graft them into new mollusks. Secondary sporocysts revert to primary sporocysts  
88 once grafted, and then undergo the normal development until cercariae emission (Jourdane &  
89 Theron 1980; Jourdane 1984).

90  
91 Here, we used a parasite strain with known genetic background (*SmBre* from Brazil; Clément *et al.*  
92 2013) to infect a sympatric *Biomphalaria glabrata* strain (*BgBre* from Brazil) and an allopatric one  
93 (*BgGua* from Guadeloupe). Interactions between *SmBre* and the two mollusk strains have been  
94 previously characterized in great detail and the differences in life traits are known (Lepesant *et al.*  
95 2013; Theron *et al.* 2014). Prevalence in the sympatric interaction (*SmBre* on *BgBre*) is 50-100%,  
96 and 30-85% in the allopatric interaction (*SmBre* on *BgGua*), depending on the number of miracidia  
97 used for infection (Theron *et al.* 2014). This shift in the environment does not represent a strong  
98 stress but has still measurable effects on the life history traits of the parasite (Lepesant *et al.*  
99 2013). We sequenced the genomes of cercariae and adult parasites to search for somatic  
100 mutations, copy number variations and mobilization of transposable elements. DNA methylation is  
101 only present at a very low level in *S. mansoni* (Geyer *et al.* 2011), which allowed us to focused on  
102 chromatin marks based on histone modifications. We used chromatin immunoprecipitation  
103 followed by sequencing (ChIP-Seq). ChIP-Seq was done on cercariae emitted after grafts, but also  
104 on adult worms obtained by infecting mice with the same cercariae. The objective was to see if  
105 chromatin structure changes, potentially induced by development in different mollusks, are  
106 transmitted to the next developmental stage of the parasite. We selected four histone  
107 modifications (H3K4me3, H3K27ac, H3K27me3, H4K20me1), associated with different biological  
108 functions. H3K4me3 form narrow peaks (rarely larger than 2 kb wide) usually associated with the  
109 transcription start site (TSS) and promoters of transcriptionally competent genes (Berger 2007; Gu  
110 & Fire 2009; Kharchenko *et al.* 2010; Zentner & Henikoff 2013). H3K27me3 plays an important role  
111 in sequential expression and regulation of developmental genes, and is generally associated with  
112 repression of transcription (Kouzarides 2007; Trojer & Reinberg 2007; Bannister & Kouzarides  
113 2011). Acetylation of H3K27 serves to discriminate active enhancers (where the mark is present)  
114 from poised ones (where the mark is absent) (Tie *et al.* 2009; Creyghton *et al.* 2010; Liu *et al.* 2011;  
115 Shlyueva *et al.* 2014). It is described as a very punctate mark, similar to H3K4me3. The role of  
116 H4K20me1 is more controversial. Some authors report it being associated with distinct regions of  
117 silent chromatin (Sims *et al.* 2006; Vielle *et al.* 2012) in human and *C. elegans*, while others  
118 describe it as an essential mark for transcription activation (Wang *et al.* 2008; Li *et al.* 2011). It is

119 also involved in DNA repair and cell cycle (Jorgensen *et al.* 2013), and genome-wide increase of  
120 this mark impairs genome integrity (Schotta *et al.* 2008).

121

122 Our data show that changes in SNP and mobilization of TE do not contribute to genetic variation,  
123 but that CNV are relatively frequent. Roughly 10% of spontaneous CNV are mitotically heritable. In  
124 addition, we observed the occurrence of environment-directed (*i.e.* environment-dependent) and  
125 genotype-independent epimutations, as well as spontaneous epimutations (not related to  
126 experimental conditions or genotypes). However, while mitotic heritability of spontaneous  
127 epimutations can reach 45%, none of the environmentally targeted chromatin structure changes  
128 are transmitted to the adult worms. Our findings are therefore in favor of spontaneous  
129 epimutations as a possible source for heritable phenotypic variants that could be selected by the  
130 environment.

131

## 132 2. Materials and Method

133

134 The experimental procedure is schematically summarized in **Figure 1**.

### 135 *Ethics statement*

136 Housing, feeding and animal care followed the national ethical standards established in the writ of  
137 February 1<sup>st</sup>, 2013 (NOR: AGRG1238753A) setting the conditions for approval, planning and  
138 operation of establishments, breeders and suppliers of animals used for scientific purposes and  
139 controls. The French Ministère de l'Agriculture et de la Pêche and French Ministère de l'Éducation  
140 Nationale de la Recherche et de la Technologie provided permit A 66040 to our laboratory for  
141 experiments on animals and certificate for animal experimentation (authorization 007083, decree  
142 87–848) for the experimenters.

### 143 *Origin of the parasites and their hosts*

144 The *Schistosoma mansoni* strain *SmBre* was originally sampled in Recife, Brazil in the 1960's. It was  
145 provided to our laboratory in 1975 by Pr. Y. Golvan (Faculté de Médecine de Paris – Saint Antoine)  
146 and has been since then maintained in its sympatric intermediate host strain *BgBre* of the mollusk  
147 *Biomphalaria glabrata* and *Mus musculus* (SWISS OF1) as definitive vertebrate host. *BgBre* (albino  
148 strain) is also from Recife and was also acquired in 1975. Albinism is of genetic origin and does not  
149 have an impact on the experimental design. The allopatric mollusk strain *BgGua* originates from  
150 the town of Dans Fond and arrived at our laboratory in 2005 (pers. comm. A. Théron). Mollusks  
151 were maintained as described in Boissier & Moné (2000).

### 152 *Infection and sex determination*

153 Thirty-two *B. glabrata* *BgBre* individuals were each infected with a single *SmBre* miracidium.  
154 Miracidia were freshly hatched from the liver of a single infected mouse. Miracidia hatching, as  
155 well as snails and mice infections, were performed as described in Boissier & Moné (2000).  
156 Mollusks were screened 30 days later for presence of parasites and cercariae shed by each  
157 positive *B. glabrata* were genotyped with sex markers (Portela *et al.* 2010). We chose to work only  
158 with female parasites, as it was previously shown that there is an increase in female cercariae  
159 production when *SmBre* develops in *BgGua* hosts (Lepesant *et al.* 2013), and also to avoid possible

160 bias caused by gender-specific epialleles. We selected two infected snails with a large number of  
161 sporocysts to be donors for sporocyst grafts. We named the female genotypes of the parasites in  
162 these two snails *Fa* and *Fb*.

### 163 *Sporocyst graft*

164 The grafts were done as detailed in Jourdane & Theron (1980) and Jourdane (1984). Large (10-12  
165 mm) *BgBre* and *BgGua* mollusks were selected as receivers. They were anesthetized by incubating  
166 4 hours in spring water mixed with sodium pentobarbital at a final concentration of 1.2 mg/mL.  
167 Donor mollusks had their shells carefully removed with tweezers and their digestive gland (where  
168 secondary sporocysts are located) was recovered. Using tweezers, explants of 1 mm<sup>3</sup> containing 1-  
169 3 sporocysts were prepared and grafted within an hour. Explants were grafted in the cephalopedal  
170 sinus of receiver snails with a custom-made glass micro-needle attached to a 1 mL syringe. A small  
171 incision of the tegument above the genital pore was done to reach the cephalopedal sinus.  
172 Grafted snails were then kept in normal growing conditions. *Fa* sporocysts were grafted in 15  
173 *BgGua* and 14 *BgBre*, and *Fb* sporocysts were transplanted in 17 *BgGua* and 15 *BgBre*.

### 174 *Parasite preparation*

175 Fifty days after transplantations, mollusks were screened under a stereomicroscope to look for  
176 secondary sporocysts. Uninfected mollusks were discarded. Cercariae were collected from  
177 infected snails 53 days after infection by gently pipetting from spring water, avoiding mucus and  
178 feces, and sedimented on ice. Water was removed and cercariae from each snail were stored  
179 separately at -80°C before being used for ChIP-Seq. As we were interested to see if epimutations  
180 induced by the development in different snail strains would be maintained in adults, we also  
181 infected mice with cercariae from some of the snails. For both *Fa* and *Fb* parasite genotype, we  
182 selected *BgBre* and *BgGua* snails emitting the highest amount of cercariae (> 5,000) to use for  
183 infection. For parasite genotype *Fa*, we had one *BgBre* and one *BgGua*, and for *Fb* we had two  
184 snails of each strain. For each snail, three mice were infected with 300 cercariae each, as described  
185 in Boissier & Moné (2000). Mice were sacrificed 42 days later by injection of sodium pentobarbital  
186 and female adult worms were recovered by retrograde perfusions of the hepatic portal system  
187 with citrate (7.5%) saline (8.5%) solution administrated through the left ventricle (Duvall & DeWitt  
188 1967). Worms trapped in the liver or mesenteric system were collected after excising these  
189 organs. Worms from each mouse were collected separately and stored at -80°C.

190

### 191 *Genomic DNA sequencing*

192 Cercariae emitted from a mollusk infected with a single miracidium are generally considered  
193 genetically identical, but evidences exist that some genetic diversity can occur (Bayne &  
194 Grevelding 2003). We could not exclude that sporocyst grafts and development in an allopatric  
195 host were source of stress, which could activate mobilization of TE (Wijayawardena *et al.* 2015).  
196 Knowing this, we performed DNA pool sequencing (Pool-Seq) on adult worms developed from  
197 cercariae emitted after the grafts. We chose to sequence DNA of adults rather than cercariae  
198 because current tools for the analysis of Pool-Seq data do not perform well on large numbers of  
199 potentially different haploid genomes. It is not feasible to obtain sufficient amount of DNA for  
200 sequencing from less than 1,000 cercariae while 10 adults are enough. We selected four pools of

201 12, 9, 11 and 17 Fb genotype female adult worms. The first two pools originated from two  
202 different sympatric (*BgBre*) snails, while the last two are from a single allopatric (*BgGua*) snail, but  
203 isolated from two different mice. DNA was extracted following Beltran *et al.* (2008) and  
204 sequencing was done as paired-end 125 bp reads on an Illumina HiSeq 2500 at the Montpellier  
205 GenomiX facility (<http://mgx.cnrs.fr>). Sequencing was designed to have a genomic coverage of at  
206 least 3X per individual per sample.

### 207 *Chromatin immunoprecipitation and sequencing (ChIP-Seq)*

208 Native ChIP was performed following the protocol developed for *Schistosoma mansoni* by Cosseau  
209 *et al.* (2009) (also available online at [http://methdb.univ-perp.fr/cgrunau/methods/native](http://methdb.univ-perp.fr/cgrunau/methods/native_chip_sm.html)  
210 [chip\\_sm.html](http://methdb.univ-perp.fr/cgrunau/methods/native_chip_sm.html)) on cercariae and adult worms. We used at least 5,000 cercariae or 20 adult worms.  
211 Immunoprecipitation was performed using the following antibodies: H3K4me3 (Millipore, cat# 04-  
212 745 lot# NG1680351, 4 µL per reaction), H3K27ac (Abcam, cat# ab4729 lot# GR-150-367-2, 4 µL  
213 per reaction), H3K7me3 (Diagenode, cat# C15410069 lot# A1821D/2, 8 µL per reaction), and  
214 H4K20me1 (Abcam, cat# ab9051 lot# GR158874-1, 4 µL per reaction). For each sample, we used a  
215 control without antibody to assess unspecific background (bound fraction) and input (unbound  
216 fraction). Inputs were used for normalization in all subsequent bioinformatics analyses. Antibodies  
217 were carefully tested for specificity as described in Cosseau & Grunau (2011) and saturating  
218 quantities were used. We had earlier shown that the abovementioned catalogue/lot numbers for  
219 anti-H3K4me3 and anti-H3K27me3 can be used for ChIP-Seq (Roquis *et al.* 2015). For H3K27ac and  
220 H4K20me1, we performed all the validation steps described by Cosseau & Grunau (2011),  
221 including a chromatin titration to assess specificity of the antibodies. Titrations, using increasing  
222 amount of antibodies over the same quantity of *S. mansoni* chromatin, proved that antibodies  
223 were indeed specific and able to immunoprecipitate all target chromatin using 4 µl of the antibody  
224 for 5,000 cercariae or 20 adult female worms (**Supplementary Figure 2**).

225 ChIP products were sequenced as single-end 50 bp reads on an Illumina HiSeq 2500 at the McGill  
226 University and Génome Québec Innovation Centre (<http://gginnovationcenter.com/index.aspx>).  
227 Briefly, fragmented DNA was quantified using a 2100 Bioanalyzer (Agilent Technologies). Libraries  
228 were generated robotically with 10 ng of fragmented DNA (range 100-300 bp) using the Kapa HTP  
229 Library Preparation Kit (Kapa Biosystems) as per the manufacturer's recommendations, except  
230 that adapters and PCR primers were diluted 100-fold. The size selection step was done after the  
231 PCR step and the number of PCR cycles was increased by six. Adapters and PCR primers were  
232 purchased from Integrated DNA Technologies whereas size selection has been performed on a  
233 Pippin Prep instrument (SAGE Biosciences Inc). Libraries were quantified using the Quant-iT  
234 PicoGreen dsDNA Assay Kit (Life Technologies) and the Kapa Illumina GA with Revised Primers-  
235 SYBR Fast Universal kit (D-Mark). Average size fragment was determined using a LabChip GX  
236 (PerkinElmer) instrument. Cluster formation on the flow cell was performed using the cBot  
237 instrument (Illumina) with four indexed libraries per lane. Sequencing, in the form of 50-cycle  
238 single end reads, was performed on a HiSeq 2000/2500 (Illumina) running HCS software version  
239 2.2.38. Demultiplexed FASTQ files were generated by allowing up to one mismatch in the index.

### 240 *Quality control, alignment and peak calling*

241 All data treatment was carried out under a local galaxy instance (Goecks *et al.* 2010)

242 (<http://bioinfo.univ-perp.fr>). Read quality was verified using the FASTX-toolkit  
243 ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)). All samples had a sanger quality score above 30 for  
244 over 95% of their length.

245 Sequences were aligned to the *S. mansoni* reference genome (Protasio *et al.* 2012) with Bowtie2.1  
246 (Langmead & Salzberg 2012) using parameters --end-to-end, --sensitive, --gbar 4 and filtered for  
247 unique match with samtools (Li *et al.* 2009) (samtools view -Sh -q *quality value 40-42* -F 0x0004 - |  
248 grep -v XS:i). SAM alignment files were converted into the bed format with pyicos (Althammer *et al.*  
249 *et al.* 2011) and sorted with sortBed -i of the bedtools suite (Quinlan & Hall 2010). PCR duplicates  
250 were removed using the remduplicates function from Pyicos (Althammer *et al.* 2011), allowing a  
251 maximum of 4 duplicates per position. Random sampling using a custom Perl script was performed  
252 on each library to generate an equal number of reads per library for peak calling. Although  
253 random sampling is not mandatory, we found in previous work that it reduces false positives when  
254 searching for histone modification differences between samples. We took 7 million reads for  
255 H3K4me3, 23 millions for H3K27ac and H3K27me3, and 19 millions for H4K20me1. One of the  
256 samples, Cercariae Fb Gua 8, had a higher amount of PCR duplicates than the other samples, and  
257 we could only select 14 million reads for H3K27ac and 10 millions for H4K20me3. Ranger of  
258 Peakranger v1.16 (Feng *et al.* 2011) with P-value cutoff 0.0001, FDR cutoff 0.01, Read extension  
259 length 200, Smoothing bandwidth 99 and Delta 1 was used for peak identification. Input samples  
260 were used as negative control (-c). Wiggle files were visualized under IGV (Thorvaldsdottir *et al.*  
261 2013). Reads mapping to repetitive regions were aligned on the 3,145 repetitive consensus  
262 sequences obtained from the same strain of *S. mansoni* (*SmBre*) in a previous study (Lepesant *et al.*  
263 *et al.* 2012) (available at <http://methdb.univ-perp.fr/downloads/>) using Bowtie2.1 (Langmead &  
264 Salzberg 2012) evoking parameters --end-to-end, --sensitive, --gbar 4. In order to obtain identical  
265 effective library sizes for each sample, the same number of aligned reads was randomly sampled  
266 for each of the ChIP experiments (H3K4me3 = 2.5 million, H3K27ac = 7 million, H3K27me3 = 7  
267 million, H4K20me1 = 8 million). After the counts had been normalized, they were compared  
268 between groups (allopatric vs sympatric, with three replicates; genotype Fa vs Fb, with two and  
269 four replicates) and between individual samples using the DESeq package (Anders & Huber 2010)  
270 of the R software. This package uses the negative binomial distribution to calculate the P-value for  
271 each element and for each comparison between two samples and to find elements that present  
272 significant differences in their enrichment levels.

### 273 *Chromatin structure profiles*

274 The analysis of the different histone modification combinations on the *S. mansoni* genome was  
275 performed using EpiCseg (Mammana & Chung 2015), using count normalization (*normalizecounts*)  
276 and binsize of 500bp (--binsize 500). We tested 4 to 16 different possible chromatin states and  
277 concluded that five states (--nstates 5) was the optimal number to avoid redundancy between  
278 states. Profiles were done on cercariae and adults, each time using the three datasets from the  
279 sympatric (BRE) condition as replicates for the bin count. In a similar way, we also compared  
280 sympatric (BRE) and allopatric (GUA) chromatin profiles for adults and cercariae. Annotation files  
281 used to generate chromatin profiles on genes, transcription start sites (TSS) and transcription end  
282 site (TES) were derived from an earlier established transcriptome (Roquis *et al.* 2015), by keeping a  
283 subset of genes longer than 5 kb (2,428 genes).



## 284 **Comparative epigenomic analysis**

285 For comparison of chromatin profiles of each mark, DiffBind v1.8.3 (Stark & Brown 2013) was used  
286 with FDR of 0.1. Comparisons were done for each histone modification between allopatric and  
287 sympatric conditions (3 replicates for each histone modification). Comparisons were also done for  
288 genotype Fa (two replicates) vs Fb (four replicates), and for each individual sample (no replicate).  
289 One library (cercariae Fb5 H3K4me3) was systematically excluded from these DiffBind analyses  
290 because it had more than 20% of PCR duplicates, indicating probably a too few starting material,  
291 thus reducing the number of replicates for this histone modification by one. Heatmaps were  
292 generated evoking dba.plot Heatmap (sample, contrast=1, correlations=FALSE) and PCA was done  
293 with dba.plotPCA(sample, contrast=1,th=.05). DiffBind (allopatric vs sympatric condition and  
294 genotype Fa vs Fb with minOverlap = 2) will test a peak if it is identified in at least two of the six  
295 total samples. It is the variance amongst all the replicates in each condition that drive the p-value  
296 calculated for the comparison. For sample-by-sample comparison, non-parametric bootstrapping  
297 or default value minOverlap = 1 was used for each sample. Both methods provided the same  
298 results. Differentially modified chromatin regions were saved as BED files for further analysis.  
299 Gene ontology enrichment analysis was done with BLAST2GO (Conesa *et al.* 2005). For analyses of  
300 significant differences in histone isoform enrichment in repetitive sequences, normalized aligned  
301 read counts were compared using the DESeq package of the R software (Anders & Huber 2010).

## 302 **Comparative genomic analysis**

303 For the detection of SNP frequency *distortion*, *i.e.* significant differences in allele frequency  
304 between sympatric and allopatric conditions, we used paired-end data for genotype Fb and single-  
305 end data corresponding to the ChIP-Seq input libraries for both Fa and Fb. All raw data were of  
306 excellent quality and no quality filter was necessary. Reads were aligned with Bowtie 2 (Langmead  
307 & Salzberg 2012) and default parameters. Unique reads were identified using the Bowtie 2 'XS:I'  
308 flag and sorted with SamTools (Li *et al.* 2009). Alignment rates were between 54% and 59%, which  
309 is consistent to what has been previously achieved on *S. mansoni* (Roquis *et al.* 2015) and can be  
310 explained by the fact that we only align on unique sequences of the genome (47% of the genome  
311 is composed of repetitive elements (Lepesant *et al.* 2012). To homogenize effective library size, 56  
312 million aligned reads were randomly subsampled for the paired-end libraries and used for further  
313 analysis. For single-end data, 24 million reads were subsampled. Pileup files were generated by  
314 SamTools invoking -B -C 0 -d 250 -q 0 -Q 13. For the comparison of allele frequencies we used  
315 PoPoolation2 (Kofler *et al.* 2011). Fisher's Exact Test was chosen to estimate the significance of  
316 allele frequency differences in samples for which we had no replicates (Fa), and the Cochran-  
317 Mantel-Haenszel (CMH) test was used to detect consistent allele frequency changes with  
318 replicates (Fb). Pileup files were converted to sync file with parameter --min-qual 20. And Fisher  
319 and CMH tests were used invoking --min-count 2 --min-coverage 10 --max-coverage 200 --  
320 suppress-noninformative for single-end data and --min-count 12 for paired-end data. Copy-  
321 number variation (CNV) are another source of genetic variability and we used CNV-seq (Xie &  
322 Tammi 2009) to estimate the probability of CNV differences between the sympatric and the  
323 allopatric conditions. We used default parameters (log2 = 0.6, pvalue = 0.001, minw = 4) with a  
324 genome size of 364,541,798 bp. Results were converted into IGV or GWAS compatible formats and  
325 visualized under IGV. Differences in genomes Fa and Fb were detected using FreeBayes v0.9.20

326 (Garrison & Marth 2012) using parameters --theta 0.001 --ploidy 2 -J -m 1 -q 0 -R 0 -Y 0 -e 1000 -F  
327 0.2 -C 2 --min-alternate-qsum 0 -G 1 --min-coverage 20. Resulting VCF files were compared with  
328 vcf-isec of the VCF tools (Danecek *et al.* 2011). CNV were searched for as described above.

329

330 The search for transposable element (TE) insertions was done by aligning the reads of the 4 Fb  
331 paired-end pools using bwa-mem from the BWA software (Li & Durbin 2009) to the reference  
332 genome of *S. mansoni*. Detection of the insertions and excision of elements was done using the  
333 TEMP software (Zhuang *et al.* 2014) with the required repeat and annotation files, parameters -x  
334 10 and -c 12 for the insertion detection, and default parameters for the absence (excision)  
335 detection. The principle of this analysis is to use differences in mapping of forward and reverse  
336 reads to detect TE insertion events. BED files were generated from the output of TEMP and the  
337 files were sorted. In order to eliminate constitutive differences between the sequenced genomes  
338 and the reference genome we subtracted the positions that were common to all four genomes  
339 and visually inspected the remaining regions.

340

### 341 *Comparative Transcriptomics*

342 Parasites were used to infect sympatric and allopatric snail hosts for three generations as  
343 described in Lepesant *et al.* (2013). For RNA extraction, cDNA generation and sequencing, and  
344 comparative analysis with the TopHat2 and CuffLinks we followed Roquis *et al.* (2015). Briefly,  
345 reads were aligned to the genome and exon-intron structures are established individually for each  
346 sample. Cuffmerge was used to generate a consensus transcription annotation file in GFF format.  
347 We then used htseq-count with the “union” option (Anders *et al.* 2015) to obtain read-counts for  
348 each of the 18,844 loci (“genes”) identified by Cufflinks/Cuffmerge. A single count  
349 (“pseudocount”) was added to each readcount to avoid division by 0 problems. DESeq (Anders &  
350 Huber 2010) with default parameters was used for differential gene expression analysis. A gene  
351 ontology enrichment analysis was done with BLAST2GO (Conesa *et al.* 2005) and prediction of  
352 protein interaction networks was done with STRING10 (Szklarczyk *et al.* 2015).

353

## 354 **Results**

### 355 *There are no significant differences in SNP or TE mobilization between SmBre in sympatric and* 356 *allopatric condition*

357 SNP frequencies were evaluated using Fisher’s exact test (for Fa without biological replicates) and  
358 the CMH test for Fb (with replicates). Both tests are robust but the CMH-test showed the highest  
359 power to identify interesting loci (Kofler & Schlotterer 2014). Following others (Barsh *et al.* 2012,  
360 Panagiotou *et al.* 2012, Kofler & Schlotterer 2014, Broer *et al.* 2013) we decided to establish a  
361 significant threshold for p-values at  $p = 1 \times 10^{-5}$ . Despite this relatively relaxed threshold, we detect  
362 no significant allele frequency distortion. We noted nevertheless that there are large blocks of 100  
363 kb range in the genome (roughly one per chromosome) that show enrichment of non-significant  
364 AF differences between the sympatric and allopatric condition in all replicates, whatever genotype  
365 was used. We calculated nucleotide diversity between Fa and Fb using private SNP (*i.e.* with allele  
366 frequency 1). We found 1,570 SNP differences (nucleotide diversity  $4.3 \times 10^{-6}$  based on the total

367 genome length of 364,541,798 bp) and 72 CNV between cercariae Fa and Fb emitted from the  
368 sympatric host *BgBre*.

369

370 The analysis of transposable element insertion produced several thousand candidate regions  
371 which showed differences in the number of supporting reads or the detection of one of the two  
372 insertion borders with respect to the reference sequence. The elimination of the redundant  
373 positions and those pointing to opposite insertion borders showed that the detected positions  
374 were due to the difference between the sequenced genome and the reference genome. Once  
375 these differences were eliminated the analysis showed that no neoinsertions of elements could be  
376 detected.

377

### 378 *There are significant differences in CNV between sympatric and allopatric SmBre*

379

380 CNV-seq detected 58 - 105 CNV covering 0.1-0.2% of the genome in all comparisons (**Table 1a &**  
381 **1b, Supplementary File 1**). Interestingly, some CNV covered or flanked the above-mentioned  
382 blocks of non-significant AF differences. To answer the question whether mitotically heritable CNV  
383 are systematically found in the allopatric condition, we compared these variants between Fa and  
384 Fb. CNV between sympatric Fa and Fb were excluded since they could represent simply differences  
385 to the reference genome. None of the regions *sensu stricto* fulfilled the condition to be a CNV in  
386 cercariae allopatric vs sympatric in both Fa and Fb, and CNV in adults allopatric vs sympatric in  
387 both Fa and Fb. Nevertheless, there is in one region (Schisto\_mansoni.Chr\_3:1,725,500-1,762,000)  
388 with a clustering of environment dependent CNV. Three single exon genes are predicted in the  
389 region (Smp\_192030, Smp\_101720 and Smp\_201970) for which we did, however, not find  
390 transcripts in our RNA-Seq data. CNV between biological replicates in sympatric or allopatric  
391 condition do not show significant differences (**Table 1a & 1b**). CNV locations are provided in the  
392 supplementary material (**Supplementary File 1**).

393

394 Taken together, our data provide evidence that, upon exposure to the allopatric condition, point  
395 mutations and TE mobilization are below the detection limit. CNV occurs at appreciable frequency,  
396 but are not linked to a specific environmental condition (allopatric or sympatric). The frequency of  
397 CNV is not different between the sympatric and allopatric hosts. Twelve to fifteen percent of  
398 random CNV are maintained through development to adults. One region on chromosome 3 could  
399 be a CNV hot-spot activated upon exposure to the allopatric environment.

400 In two cases (Schisto\_mansoni.Chr\_2:717900-817000 and Schisto\_mansoni.Chr\_1:58049000-  
401 58059000), we found epimutations close to CNV, but not to SNP or transposable elements. After  
402 having established the degree and type of genetic mutations we attempted to characterize the  
403 epimutations.

### 404 *The epigenome of S. mansoni comes in five main chromatin colors*

405 We then performed CHIP-Seq experiments with the abovementioned antibodies for two  
406 genotypes (Fa and Fb) of cercariae that had developed in sympatric or allopatric snail hosts, and of  
407 the resulting female adults (a total of 60 libraries, details in **Supplementary File 2**). The analysis of  
408 the different histone modification combinations on the *S. mansoni* genome shows five main states

409 (also called “colors”; Filion *et al.* 2010), in both cercariae (**Figure 2**) and adults (**Supplementary**  
410 **Figure 3**). Both profiles are qualitatively similar. The most frequent histone mark combination  
411 (state 3 on **Figure 2** and **Supplementary Figure 3**) is a combination of H3K27me3-H3K27ac-  
412 H4K20me1, which covers 60% of the genome in cercariae and only 46% in adults. This state has a  
413 low presence around the transcription start site of genes (TSS) but is strongly enriched at the  
414 transcription end site (TES). The second most frequent state (state 2) is characterized by an  
415 enrichment of H3K27ac and H3K20me1, covering 17% of the genome of cercariae and 24% of  
416 adults. It is enriched upstream of the TSS, decreases drastically in the first kb past the TSS before  
417 being enriched again in the gene body and decreases once more near the 3’end of the gene. The  
418 absence of any of the four histone marks (state 5) covers 14% of the genome in both  
419 developmental stages of the parasite. It is not very common in gene bodies, but is particularly  
420 depleted around TSS and TES. State 4, characterized by a stronger presence of H3K27me3,  
421 combined with a lower amount of H3K27ac and H4K20me1 is very lowly represented around  
422 genes. Outside of genic regions, it can be found covering 5% of cercariae and 14% of adult  
423 genomes. The last combination of marks (state 1) presents a very strong enrichment in H3K4me3  
424 located right after the TSS (3% of cercariae and 2.3% of adult genomes). We also compared  
425 chromatin state between sympatric and allopatric cercariae, as well as adults, but we did not  
426 observe any large-scale chromatin structure profile differences (data not shown). It should be  
427 noted that except for H3K27ac and H3K27me3, that are mutually exclusive, our approach can  
428 neither confirm nor exclude that histone modifications occur at the same histone or same  
429 nucleosome. All we can say it that there is co-occurrence at the same genomic site.

430

### 431 *There are three types of chromatin changes upon exposure to allopatric hosts*

432 For comparison of chromatin changes in the sympatric and allopatric conditions, one would ideally  
433 use these five chromatin colors. Unfortunately, no algorithm exists yet to do so. Instead, we  
434 performed what was feasible, *i.e.* a comparative analysis histone mark by histone mark for the  
435 four modifications using peak-calling (PeakRanger) and differential enrichment analysis (DiffBind).  
436 For the cercaria we found in average 8,378 peaks for H3K4me3, 15,866 peaks for H3K27ac, 20,784  
437 peaks for H3K27me3, and 11,175 peaks for H4K20me1. In adults there were 11,967 peaks for  
438 H3K4me3, 9,1110 peaks for H3K27ac, 22,435 peaks for H3K27me3, and 17,403 peaks for  
439 H4K20me1 (**Supplementary File 3**).

440

441 Heatmaps generated by DiffBind provide a general overview of the chromatin landscape. They  
442 clearly show chromatin differences in cercariae clones exposed to the sympatric and allopatric  
443 host, and the resulting female adults (**Figure 3A** and **3B**). Principal component analysis, using  
444 genomic positions with differences on chromatin profiles, separated sympatric and allopatric  
445 samples extremely well, based on the number of differences in histone marks between samples  
446 (**Supplementary Figure 4A** and **4B**). We observed in cercariae three different types of chromatin  
447 changes depending on condition and genotype. In principle, to classify epimutations, one would fit  
448 a statistical model to test for a genotype effect, an environmental effect, and possibly the  
449 interaction, on peak presence and shape. In such cases one could detect a significant effect of (for  
450 instance) environment even if not all replicates showed the pattern consistently. To our

451 knowledge, such statistical model has not been developed and tested. We use here a simpler  
452 approach and classified epimutations into **(1)** host-induced epimutations, **(2)** spontaneous  
453 epimutations, and **(3)** genotype dependent epimutations. For convenience, we will also use the  
454 term “targeted epimutation” or “induced epimutations” for type 1 epimutation, and “random  
455 epimutations” for type 2. In other words, at a given developmental stage (cercariae or adults),  
456 type 1 epimutations must be present in all samples of the allopatric condition and absent in the  
457 sympatric (or vice versa). Type 2 epimutations are present in at least one sample of a condition or  
458 a genotype, but must not be present in all samples of this condition or genotype. Type 3  
459 epimutations are present in all samples of one genotype (independently of the environmental  
460 condition) and absent in all samples of the other genotype. A potential caveat of this classification  
461 method lies in the relatively low number of replicates, leading potentially to an overestimation of  
462 type 1 epimutations. In addition, type 2 epimutations might not be stochastic in the strict sense.  
463 For instance, they might also be environment or genotype-specific but with incomplete  
464 “penetrance” among replicate samples.

465  
466 For type 1 epimutations, using a FDR 0.1 (Diffbind’s default), we identified 131 differences for  
467 H4K20me2, 592 for H3K4me3, 346 for H3K27ac and 203 for H3K27me3, present in all replicates  
468 between sympatric and allopatric cercariae. All were verified by visual inspection, in which we  
469 looked for background changes that DiffBind had probably mistakenly identified as real  
470 differences. Basically, if a trained human eye (considering intensity, shape, and strength of  
471 neighboring peaks) cannot detect the differences spotted by Diffbind, we considered it as a false  
472 positive (see **Supplementary Figure 5** for examples). In general, validated epimutations  
473 corresponded to a p-value  $\leq 1 \times 10^{-5}$  for H3K4me3 and  $\leq 1 \times 10^{-15}$  for the other marks. It reduced the  
474 number of type 1 differences to a total of 51 genomic regions, among those all of them showed  
475 changes for H3K4me3, 11 for H3K27ac, 2 for H3K27me3, and 3 for H4K20me1. Examples are  
476 shown in **Figure 4A** and a detailed list of differences is available in **Supplementary File 4**.  
477 Interestingly, epigenetic diversity is higher in samples from the allopatric condition (**Figure 3A** and  
478 **Supplementary Figure 4A**). We found that 76,5% of induced epimutations are in transcribed  
479 regions previously identified with RNA-Seq in cercariae (Roquis *et al.* 2015).

480  
481 As outlined above we defined spontaneous epimutations (type 2) as a chromatin structure change  
482 that occurs in one or more samples, but that is not consistently shared among replicates (**Figure**  
483 **4B**). We detected 64 genomic locations where these random epimutations occurred, very often  
484 with multiple changes in the same region: 60 changes in H3K4me3, 9 in H3K27ac, 6 in H3K27me3,  
485 and 5 in H4K20me1. We identified that 79.7% of the differences were in transcribed regions  
486 (details available in **Supplementary File 4**). We detected these epimutations by comparing  
487 samples one by one on Diffbind results (**Supplementary File 5**), and verifying them through the  
488 same visual inspection as for type 1 epimutations. We noted that for the two replicates of  
489 genotype Fb in the allopatric condition, we systematically found several hundreds of random  
490 epimutation for the four histone isoforms. We believe that this is due to a technical issue leading  
491 to higher background in these ChIP-Seq libraries, which is consistent with our visual inspection.

492

493 We also looked for genotype related epimutations (type 3), *i.e.* chromatin structure changes that  
494 are identical in both environmental conditions, but different when comparing cercariae of  
495 genotype Fa and Fb (**Figure 4C**). After visual inspection (same as type 1 & 2), we observed a total  
496 of 64 genomic locations with chromatin structure differences between the two genotypes, often  
497 with two or more marks changing: 39 differences for H4K20me2, 41 for H3K4me3, 38 for H3K27ac  
498 and 42 for H3K27me3. 73.4% of the changes between Fa and Fb are in transcribed genomic  
499 regions. BED files are available as **Supplementary File 5**.

500

### 501 *Spontaneous epimutations, but not host-induced epimutations in cercariae are inherited to* 502 *adults*

503 By infecting mice with some of the cercariae collected from each sympatric and allopatric snail, we  
504 produced adult worms and investigated if the different types of epimutations we had spotted in  
505 cercariae were transmitted to the next developmental stage. None of the type 1 chromatin  
506 structure changes, shared by all replicates, were maintained in adults. However, Diffbind  
507 heatmaps (with the exception of H3K4me3) still show differences between adults developed from  
508 cercariae that have experienced either the sympatric or allopatric condition (**Figure 3B**), regardless  
509 of the genotype (Fa or Fb). With the exception of H3K27me3, PCA of adults separate adults by the  
510 environmental condition they experienced as cercariae (**Supplementary Figure 4A**). While host-  
511 induced epimutations were not inherited *sensu stricto* from cercariae to adults, the developmental  
512 trajectory was altered, resulting in different chromatin profiles between adults from allopatric or  
513 sympatric origin (**Figure 3A and 3B; Supplementary Figure 4A and 4B**).

514

515 Since we had replicates for genetically identical clones Fb, we investigated whether spontaneous  
516 epimutation (type 2) would be transmitted to the corresponding adults. Indeed, 45% of the sites  
517 with epimutations have the same chromatin structure in adults: 29 differences in H3K4me3, 3 in  
518 H3K27ac, 5 in H3K27me3 but none in H4K20me1 were heritable *sensu stricto*, *i.e.* they had  
519 occurred in the sporocysts and were maintained in cercariae and adults.

520

521 We found genotype-associated epimutations to be most mitotically heritable. Roughly 77% of the  
522 genomic locations with epimutations between Fa and Fb cercariae were found to be identical in  
523 adults. It corresponds to 31 changes in H3K4me3, 23 in H3K27ac, 24 in H3K27me3 and in 30  
524 H4K20me1.

525

526 We then used the average number of peaks detected by PeakRanger (**Supplementary File 3**) to  
527 calculate frequency and heritability of histonic epimutations. Environmentally triggered  
528 epimutations (type 1) occur at a frequency of  $4.2 \times 10^{-3}$  (H3K4me3),  $1.2 \times 10^{-3}$  (H3K27ac),  $8.9 \times 10^{-5}$   
529 (H3K27me3), and  $1.7 \times 10^{-4}$  (H4K20me1). Random epimutations (type 2) are observed in the same  
530 order of magnitude:  $5 \times 10^{-3}$  (H3K4me3),  $1 \times 10^{-3}$  (H3K27ac),  $2.7 \times 10^{-4}$  (H3K27me3), and  $2.9 \times 10^{-3}$   
531 (H4K20me1). As mentioned above, heritability of type 1 epimutations is 0 (reversibility = 1). In  
532 contrast, heritability of type 2 epimutations reaches up to 0.83. In summary, frequency and  
533 heritability of histonic epimutations depends on the chemical nature of the modification and their  
534 type. Mean frequency is in the  $10^{-3}$  order of magnitude.

535

536 Based on their locations, epimutations for the different histone marks were manually compiled  
537 into a single table for further analysis (**Supplementary File 4**).  
538

#### 539 *Negligible chromatin structure changes occur around repetitive elements*

540 Mapping our ChIP-Seq library over the 3,145 identified repetitive sequences of *S. mansoni* genome  
541 and analyzing the read counts between samples, we did not find any differences present in all  
542 replicates of an environmental condition in cercariae or adults. Only when we compared  
543 environmental conditions within a single genotype, or between genotypes in the same conditions,  
544 we detected changes in 21 repetitive elements (17 of which being transposable elements). Only a  
545 single histone mark was modified in these 21 repetitive elements. None of the histone marks were  
546 systematically changed and we found no heritability of the changes from cercariae to adults  
547 (**Supplementary File 6**).  
548

#### 549 *Chromatin structure changes are not associated with specific pathways or molecular/biological 550 functions*

551 After having established that chromatin architecture changes upon exposure to a different  
552 environment and that this occurs primarily in regions of the genome that contain genes, we  
553 wondered whether a phenotype was associated with these epimutations. GO enrichment analysis  
554 was done for all the transcribed regions of the genome where chromatin structure changes were  
555 detected (host-induced, spontaneous and genotype-associated epimutations). No specific  
556 pathway or enrichment in molecular or biological functions was observed in any type of  
557 comparison. STRING analysis showed that Smp\_016630 (Trimeric G-protein alpha o subunit),  
558 Smp\_141980 (cAMP-specific 3,5-cyclic phosphodiesterase) and Smp\_147260 (Ras-like) could  
559 potentially interact in a signaling pathway. All three genes have environmentally triggered  
560 complex epimutations (**Supplementary File 4**).  
561

562 Next, we exposed miracidia for three generations to the allopatric host, extracted RNA from the  
563 resulting cercariae and compared their RNA profile to cercariae produced in sympatric conditions.  
564 We had earlier shown that transcription is stalled in cercariae (Roquis *et al.* 2015) and we  
565 expected only small differences in RNA profiles. We compared the RNA levels of genes at or close  
566 to the type 1 epimutations using the three generations as biological replicates (**Supplementary  
567 File 7**). Differences in RNA levels of epimutated genes were moderate with a few exceptions:  
568 Smp\_130050 codes probably for a PDZ domain containing RING finger protein and its RNA is 3-5  
569 fold overrepresented in allopatric cercariae. Smp\_151290 codes for a putative multidrug and toxin  
570 extrusion protein 2 and its RNA is 2.5 times enriched in cercariae from allopatric hosts.  
571 Smp\_015630 (Glutamate-gated chloride channel subunit) is threefold underrepresented in  
572 allopatric cercariae.  
573

574 We also identified genes with strong differences *i.e.* adjusted p-value < 0.01, and/or 10 fold  
575 overrepresentation in one of the conditions (**Supplementary File 8**). For these genes we verified  
576 chromatin structure in cercariae and adults. None of the genes were in the regions that we had  
577 qualified as epimutated. However, it is noteworthy that 80% (50 of 62) of these genes do not

578 contain a canonical H3K4me3 peak at the TSS. This is two times higher than the genome average.  
579 It could therefore be that with the antibody combination we used we did not capture the  
580 chromatin structure of these genes. We also noticed subtle differences in the extent and position  
581 of chromatin states 1 to 5 (**Supplementary Figure 6**). We would have considered this as random  
582 noise but in some cases the chromatin state profile was faithfully inherited to the adults,  
583 suggesting that these small differences are stable and might have an impact on gene expression.  
584

## 585 Discussion

586  
587 In this experiment we set out to investigate the impact of the environment on the genome and epigenome  
588 of *S. mansoni* larvae. The environment could be (i) a selective agent, (ii) could induce specific mutations  
589 and/or epimutations, and (iii) could change mutation and/or epimutation rate, or (iv) might have no effect.  
590 The two *S. mansoni* genotypes we used in this study (Fa and Fb) come from the same inbred  
591 population. Nucleotide diversity is 0.0004% based on SNP, but they show relatively strong  
592 differences on the level of CNV (0.2% of the genome). We also found that they bear epigenetic  
593 differences. Here, we considered genotype-associated epimutations as chromatin structure  
594 changes present within a genotype in both allopatric and sympatric condition, but absent in the  
595 other genotype. We observed 64 of these differences ( $FDR \leq 0.1$ ,  $p \leq 1 \times 10^{-5}$  and after visual  
596 inspection). The majority (73.4%) of epimutations are in transcribed regions in cercariae, which is  
597 similar to what we observed with spontaneous and induced epimutations. In at least two striking  
598 cases that we investigated, genotype associated epimutations were linked to CNV. In the case of  
599 clone Fa, there is an important CNV difference between cercariae emitted from sympatric and  
600 allopatric snails, and could be a result of the environment, although without replicates for this  
601 genotype we cannot assess it without doubt. It has been previously observed (Skinner *et al.* 2015)  
602 that environmentally induced epimutations can lead to genetic mutation and CNV. We noticed  
603 that the CNV were clustered (**Supplementary File 1**). Centromere positions on *S. mansoni* genome  
604 are not mapped to the genome, and it is also possible that these CNV clusters are located in  
605 pericentromeric regions, which are hotspots of segmental duplications and prone to generate CNV  
606 (Makino *et al.* 2013). Apparition of these CNV could also possibly explain some of the genetic  
607 heterogeneity between *S. mansoni* clonal sporocysts observed before (Bayne & Grevelding 2003).  
608

609 We then looked at environmentally triggered epigenetic changes, *i.e.* epimutations that occur in  
610 all the samples of a given mollusk condition. We found a total of 51 genomic regions, where at  
611 least one histone mark is different in all three cercariae samples emitted from sympatric and  
612 allopatric mollusks. Changes were often subtle (peaks of low intensities compared to neighbor  
613 peaks). As we used whole organisms for ChIP-Seq, there is the possibility that these changes only  
614 reflect a small subset of cells or tissues within cercariae. No gene pathway or gene ontology term  
615 enrichments were uncovered in genes spanning induced epimutations, and none of these  
616 epimutations were transmitted to adults. This type of ephemeral epigenetic change in response to  
617 an environmental stimulus is reminiscent of what was found for instance in smokers and former  
618 smokers where roughly a dozen CpG sites change their methylation status but revert progressively  
619 to normal when smoking was abandoned (Wan *et al.* 2012; Elliott *et al.* 2014).  
620



621 No major chromatin remodeling was observed in the 3,145 known repetitive sequences of the *S.*  
622 *mansoni* genome, and although some of the epigenetic differences are located in transposable  
623 elements, we did not detect new insertion sites. Transposable elements movement is known to  
624 generate new genetic, epigenetic and phenotypic diversity in other works (Biemont 2010; Fedoroff  
625 2012; Holoch & Moazed 2015), but they do not play a major role in this study.

626

627 Subsequently, we investigated the presence of spontaneous (type 2) epimutations through the  
628 genome of cercariae. We first detected type 2 epimutations through comparison of replicates of  
629 clone Fb. Sixty-four sites on the genome showed random epimutation, and 45% of them  
630 maintained their epimutated chromatin structure at the adult stage. Our data suggest that a  
631 change of the environment has an impact on epimutations since chromatin profiles of cercariae  
632 emitted from *BgGua* (and the resulting adults) present more differences between samples (as  
633 observed from the PCA on **Supplementary Figure 4A** and **4B**) than chromatin profiles of cercariae  
634 and adults emitted from *BgBre*. We had shown before that *S. mansoni* worms exposed transiently  
635 to hycanthone, an antelmintic drug, possess 57 days later a characteristic chromatin profile  
636 different from unexposed worms (Roquis *et al.* 2013). At that time, we were not able to tell if the  
637 drug directly induced the different profiles, or if it acted as a selection agent, killing most of the  
638 parasites in the population, except those with a pre-existing distinct chromatin structure. In light  
639 of our new findings we are now in favor of the hypothesis that random epimutations had occurred  
640 prior to the drug exposure, and had produced in some individuals a resistant phenotype to the  
641 treatment.

642

643 The average frequency of histone modification based epimutations in our system is about  $10^{-3}$ .  
644 This is one order of magnitude higher than what was reported for DNA methylation based  
645 epimutations in CpG pairs (Schmitz *et al.* 2011; van der Graaf *et al.* 2015). In this issue, Kronholm &  
646 Collins (2015) underline the importance of stability of epimutations for adaptive evolution. We  
647 show that heritability (*i.e.* stability) depends strongly on the type of epimutation and that for most  
648 of the detected epimutations reversion is actually above the threshold of  $10^{-1}$  that Kronholm and  
649 Collins define as the reversion rate that limits the adaptive effect of epimutations. In other words,  
650 our data support the view that in *S. mansoni* epimutations do not contribute to adaptive walks,  
651 unless they provide a strong fitness gain.

652

653 *Schistosoma mansoni* proves to be an interesting model for this kind of epigenetic studies, due to  
654 its ecological/healthcare importance, the possibility to generate clones, and different  
655 phenotypically distinct developmental stages allowing to observe mitotic inheritance of  
656 epimutations. However, one of the caveats of this organism is its small size, making it impossible  
657 at the moment to perform ChIP-Seq on specific tissues or individuals (we needed at least 5,000  
658 cercariae for our experiments). It means that we can only observe epimutations that are present  
659 in most of the cells and most individuals in the pools. Tissue-specific epimutations, as well as rare  
660 epimutations within the clone population, cannot be detected. Conversely, this also means that  
661 the chromatin structure changes we observed concern the majority of cells and individuals in our  
662 pools, and hence are more likely to have a significant adaptive impact. We also keep in mind that  
663 our sample size (three for each environmental condition and developmental stage) is small, and

664 more replicates, from the same and new genotypes, need to be done in order to strengthen our  
665 findings.

666

## 667 **Conclusion**

668 In conclusion, our findings indicate that in *S. mansoni* sporocysts, there are two major sources of  
669 mitotically heritable variation: CNV and spontaneous epimutations. This means that at least in our  
670 biological model (i) epimutations occur at appreciably frequency and since they are  
671 overrepresented in coding regions they can have an impact on gene expression, and (ii) that CNV  
672 occur at the same order of magnitude as epimutations, potentially contributing as much as the  
673 latter to phenotypic variation. We could not fully answer the question whether or not  
674 environmental changes increase epimutations. That means that at least for *S. mansoni* theoretical  
675 models in which epimutations serve to produce phenotypic variants that explore the fitness  
676 landscape for invariant genotypes (e.g. Klironomos *et al.* 2013) are fully applicable but future  
677 models must include CNV and must take into account the high reversion rate of histone-based  
678 epimutations.

679

## 680 **Acknowledgement and data availability**

681 Authors are grateful to Rory Stark, Aaron Taudt, Maria Colomé-Tatche and Frank Johannes for  
682 helpful discussions. Data for ChIP-Seq, Pool-Seq and RNA-Seq are available as fastq files BioProject  
683 Accession PRJNA236156 at the NCBI SRA. Chromatin profiles and transcripts can be visualized at  
684 our GBrowse instance (Stein *et al.* 2002) (<http://genome.univ-perp.fr>). Funding was provided by  
685 the French National Research Agency ANR, project ANR-10-BLAN-1720 (EPIGEVOL) and the GenEpi  
686 program of the UMR5244/IHPE (<http://ihpe.univ-perp.fr/en/ihpe-transversal-epigenetics/>).

687

## 688 **Figure legends**

689

690 **Figure 1:** Experimental procedure for the sporocyst grafts. Brazilian mollusks (*BgBre*) were  
691 infected with a single miracidium of the sympatric strain of *S. mansoni* (*SmBre*). Thirty days later,  
692 secondary sporocysts were surgically removed from the infected mollusks and grafted in several  
693 sympatric (*BgBre*) and guadeloupean allopatric (*BgGua*) mollusks. Emitted cercariae were  
694 collected for chromatin immunoprecipitations followed by sequencing (ChIP-Seq) and infection of  
695 mice to obtain adult worms. Adult worms were recovered for ChIP-Seq and paired-end DNA  
696 sequencing. The experiment was performed twice, with two different genotypes of the parasite  
697 (Fa and Fb). In Fb genotype, two replicates were done in sympatric and allopatric conditions.

698

699 **Figure 2:** Five different chromatin states (“colors”) as identified by EpiCSeq in cercariae. **(A)**  
700 Composition in histone marks of the five chromatin states. Color code of the states is used in all  
701 other parts of this figure. **(B)** Proportion of the genome covered by each state in cercariae.  
702 Enrichment for each state around gene bodies **(C)**, at transcription start site (TSS) **(D)** and  
703 transcription end site (TES) **(E)** of cercariae. Enrichment profiles were done on 2,428 genes longer  
704 than 5 kb. Results for adults are very similar and shown in **Supplementary Figure 2.**

705

706 **Figure 3:** Heatmaps for the four studied histone marks in cercariae (A) and adults (B). Only sites  
707 with statistically significant differences (epimutations) were used. Red bars highlight conserved  
708 changes. Red snail symbols represent sympatric host *B. glabrata* BgBre, blue snail symbols stand  
709 for allopatric snail host *BgGua*. Cercariae sample GuaFb5 for H3K4me3 was excluded from analysis  
710 (see methods).

711

712 **Figure 4:** Typical examples of epimutations. (A) environmentally dependent epimutations (green  
713 frame) in H3K4me3 (blue profile) and H3K27ac (red profile): there are differences in all three  
714 samples Fa and two replicates Fb between sympatric and allopatric cercariae. (B) Random  
715 epimutations in H3K4me3, present in cercariae in three samples and transmitted to adult worms  
716 in corresponding three samples. (C) genotype-associated epimutations in H3K4me3, H3K27ac,  
717 H3K27me3 and H4K20me1 between genotypes Fa and Fb in sympatric condition. The track Ref.  
718 transcriptome *S. mansoni* displays transcripts (exons as blocks and introns as arrows) deduced  
719 from RNA-Seq data in cercariae.

720

721 **Supplementary Figure 1:** Life cycle of the parasite *Schistoma mansoni*. The endoparasites infect  
722 two hosts, (1) a mollusk intermediate host (*Biomphalaria* genus) in which it multiplies asexually as  
723 sporocysts and then cercariae, and (2) a vertebrate definitive host (primate or rodent) in which  
724 adult worms reproduce sexually. Two types of free-swimming larvae, miracidia and cercariae, are  
725 the infective form for the mollusk and vertebrate hosts, respectively. Eggs laid by adult worm  
726 couples in the mesenteric vein of the vertebrate host can go through the intestine wall to be  
727 excreted in the feces, and hatch in contact with freshwater.

728

729 **Supplementary Figure 2:** ChIP titration, using the same amount of *S. mansoni* chromatin  
730 (equivalent to 20 adult worms) and an increasing amount of antibody. Percentage of input  
731 recovery of chromatin was measured by qPCR on two housekeeping genes, alpha tubulin and 28S,  
732 as described in (Cosseau *et al.* 2009). Titrations were done to measure the efficiency of  
733 immunoprecipitation for antibody H3K27ac Abcam cat# ab4729 lot# GR-150-367-2 (A and B) and  
734 H4K20me1 Abcam, cat# ab9051 lot# GR158874-1 (B and C).

735

736 **Supplementary Figure 3:** Five different chromatin states (“colors”) as identified by EpiCseg in  
737 adults. (A) Composition in histone marks of the five chromatin states. Color code of the states is  
738 used in all other parts of this figure. (B) Proportion of the genome covered by each state in adults.  
739 Enrichment for each state around gene bodies (C), at transcription start site (TSS) (D) and  
740 transcription end site (TES) (E) of adult worms. Enrichment profiles were done on 2,428 genes  
741 longer than 5 kb.

742

743 **Supplementary Figure 4:** Principal Component Analysis (PCA) for the four studied histone marks in  
744 cercariae (A) and adults (B). Only sites with statistically significant differences (épimutations) were  
745 used. Blue circles (Bre) represent cercariae and adults from sympatric conditions (*SmBre* in *BgBre*),  
746 magenta (Gua) the allopatric conditions (*SmBre* in *BgGua*). Cercariae sample GuaFb5 for H3K4me3  
747 was excluded from analysis (see methods).

748

749 **Supplementary Figure 5:** Examples of two regions (**A** and **B**) that were identified as epimutations  
750 (displayed by the red boxes) by DiffBind but did not pass the visual inspection.

751

752 **Supplementary Figure 6:** Example of a region that was not identified as epimutated by DiffBind  
753 but shows differences based on EpiCSeq chromatin “colors”, judged by visual inspection. The  
754 associated gene is differentially expressed and chromatin states are transmitted to adults.

755

## 756 **References**

757

758 Althammer S, Gonzalez-Vallinas J, Ballare C, Beato M, Eyraas E (2011) Pyicos: a versatile toolkit for  
759 the analysis of high-throughput sequencing data. *Bioinformatics*, **27**, 3333–3340.

760 Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome*  
761 *Biology*, **11**, R106.

762 Anders S, Pyl PT, Huber W (2015) HTSeq—a Python framework to work with high-throughput  
763 sequencing data. *Bioinformatics*, **31**, 166–169.

764 Bannister AJ, Kouzarides T (2011) Regulation of chromatin by histone modifications. *Nature*  
765 *Publishing Group*, **21**, 381–395.

766 Barsh GS, Copenhaver GP, Gibson G, Williams SM (2012) Guidelines for genome-wide association  
767 studies. *PLoS Genetics*, **8**, e1002812.

768 Bayne CJ, Grevelding CG (2003) Cloning of *Schistosoma mansoni* sporocysts in vitro and detection  
769 of genetic heterogeneity among individuals within clones. *Journal of Parasitology*, **89**, 1056–  
770 1060.

771 Beltran S, Galinier R, Allienne JF, Boissier J (2008) Cheap, rapid and efficient DNA extraction  
772 method to perform multilocus microsatellite genotyping on all *Schistosoma mansoni* stages.  
773 *Memorias do Instituto Oswaldo Cruz*, **103**, 501–503.

774 Berger SL (2007) The complex language of chromatin regulation during transcription. *Nature*, **447**,  
775 407–412.

776 Biemont C (2010) A brief history of the status of transposable elements: from Junk DNA to major  
777 players in evolution. *Genetics*, **186**, 1085–1093.

778 Boissier J, Moné H (2000) Experimental observations on the sex ratio of adult *Schistosoma*  
779 *mansoni*, with comments on the natural male bias. *Parasitology*, **121 ( Pt 4)**, 379–383.

780 Broer L, Lill C, Schuur M *et al.* (2013) Distinguishing true from false positives in genomic studies: p  
781 values. *European Journal of Epidemiology*, **28**, 131–138.

782 Clément JAJ, Toulza E, Gautier M *et al.* (2013) Private selective sweeps identified from next-  
783 generation pool-sequencing reveal convergent pathways under selection in two inbred  
784 *Schistosoma mansoni* strains (AR Jex, Ed.). *PLoS Neglected Tropical Diseases*, **7**, e2591.

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

787 Cosseau C, Grunau C (2011) Native chromatin immunoprecipitation. *Methods in molecular biology*  
788 (*Clifton, N.J.*), **791**, 195–212.

- 789 Cosseau C, Azzi A, Smith K *et al.* (2009) Native chromatin immunoprecipitation (N-ChIP) and ChIP-  
790 Seq of *Schistosoma mansoni*: Critical experimental parameters. *Molecular and Biochemical*  
791 *Parasitology*, **166**, 70–76.
- 792 Creyghton MP, Cheng AW, Welstead GG *et al.* (2010) Histone H3K27ac separates active from  
793 poised enhancers and predicts developmental state. *Proceedings of the National Academy of*  
794 *Sciences*, **107**, 21931–21936.
- 795 Danchin E, Charmantier A, Champagne FA *et al.* (2011) Beyond DNA: integrating inclusive  
796 inheritance into an extended theory of evolution. *Nature Reviews Genetics*, **12**, 475–486.
- 797 Danecek P, Auton A, Abecasis G *et al.* (2011) The variant call format and VCFtools. *Bioinformatics*,  
798 **27**, 2156–2158.
- 799 DeJong RJ, Morgan JA, Paraense WL *et al.* (2001) Evolutionary relationships and biogeography of  
800 *Biomphalaria* (*Gastropoda: Planorbidae*) with implications regarding its role as host of the  
801 human bloodfluke, *Schistosoma mansoni*. *Molecular Biology and Evolution*, **18**, 2225–2239.
- 802 Dias LC de S, Bruce JI, Coles GC (1988) Strain variation in the infectivity of *Schistosoma mansoni* for  
803 *Biomphalaria glabrata*. *Revista do Instituto de Medicina Tropical de São Paulo*, **30**, 86–90.
- 804 Duncan EJ, Gluckman PD, Dearden PK (2014) Epigenetics, plasticity, and evolution: How do we link  
805 epigenetic change to phenotype? *Journal of Experimental Zoology Part B: Molecular and*  
806 *Developmental Evolution*, **322**, 208–220.
- 807 Duvall RH, DeWitt WB (1967) An improved perfusion technique for recovering adult schistosomes  
808 from laboratory animals. *The American journal of tropical medicine and hygiene*, **16**, 483–486.
- 809 Elliott HR, Tillin T, McArdle WL *et al.* (2014) Differences in smoking associated DNA methylation  
810 patterns in South Asians and Europeans. **6**, 1–10.
- 811 Fedoroff NV (2012) Transposable elements, epigenetics, and genome evolution. *Science (New*  
812 *York, N.Y.)*, **338**, 758–767.
- 813 Feng X, Grossman R, Stein L (2011) PeakRanger: A cloud-enabled peak caller for ChIP-seq data.  
814 *BMC Bioinformatics*, **12**, 139.
- 815 Filion GJ, van Bemmel JG, Braunschweig U *et al.* (2010) Systematic protein location mapping  
816 reveals five principal chromatin types in *Drosophila* cells. *Cell*, **143**, 212–224.
- 817 Garrison E, Marth G (2012) *Haplotype-based variant detection from short-read sequencing. 2012.*
- 818 Geyer KK, pez CMRIGLO, Chalmers IW *et al.* (2011) Cytosine methylation regulates oviposition in  
819 the pathogenic blood fluke *Schistosoma mansoni*. *Nature Communications*, **2**, 424–10.
- 820 Goecks J, Nekrutenko A, Taylor J, Galaxy Team T (2010) Galaxy: a comprehensive approach for  
821 supporting accessible, reproducible, and transparent computational research in the life  
822 sciences. *Genome Biology*, **11**, R86.
- 823 Gu SG, Fire A (2009) Partitioning the *C. elegans* genome by nucleosome modification, occupancy,  
824 and positioning. *Chromosoma*, **119**, 73–87.
- 825 Guerrero-Bosagna C, Skinner MK (2011) Environmentally induced epigenetic transgenerational  
826 inheritance of phenotype and disease. *MOLECULAR AND CELLULAR ENDOCRINOLOGY*, **354**, 1–  
827 6.
- 828 Holoch D, Moazed D (2015) RNA-mediated epigenetic regulation of gene expression. *Nature*  
829 *Reviews Genetics*, **16**, 71–84.
- 830 Jablonka E, Lamb MJ (2008) Soft inheritance: challenging the modern synthesis. *Genetics and*

- 831 *Molecular Biology*, **31**, 389–395.
- 832 Jablonka E, Lachmann M, Lamb MJ (1992) Evidence, mechanisms and models for the inheritance  
833 of acquired characters. *Journal of Theoretical Biology*, **158**, 245–268.
- 834 Jiang C, Mithani A, Belfield EJ *et al.* (2014) Environmentally responsive genome-wide accumulation  
835 of de novo *Arabidopsis thaliana* mutations and epimutations. *Genome Research*, **24**, 1821–  
836 1829.
- 837 Jorgensen S, Schotta G, Sorensen CS (2013) Histone H4 Lysine 20 methylation: key player in  
838 epigenetic regulation of genomic integrity. *Nucleic Acids Research*, **41**, 2797–2806.
- 839 Jourdane J (1984) Maintenance of a male and a female clone of *Schistosoma mansoni* by  
840 microsurgical transplantation of sporocysts. Reliability of the method. *Annales de*  
841 *parasitologie humaine et comparée*, **59**, 361–367.
- 842 Jourdane J, Theron A (1980) *Schistosoma mansoni*: cloning by microsurgical transplantation of  
843 sporocysts. *Experimental parasitology*, **50**, 349–357.
- 844 Kharchenko PV, Alekseyenko AA, Schwartz YB *et al.* (2010) Comprehensive analysis of the  
845 chromatin landscape in *Drosophila melanogaster*. *Nature*, **471**, 480–485.
- 846 King CH (2010) Parasites and poverty: The case of schistosomiasis. *Acta Tropica*, **113**, 95–104.
- 847 Klironomos FD, Berg J, Collins S (2013) How epigenetic mutations can affect genetic evolution:  
848 Model and mechanism. **35**, 571–578.
- 849 Knight M, Arican-Goktas HD, Ittiprasert W *et al.* (2013) Schistosomes and snails: a molecular  
850 encounter. *Frontiers in genetics*, **5**, 230–230.
- 851 Kofler R, Schlotterer C (2014) A guide for the design of evolve and resequencing studies. *Molecular*  
852 *Biology and Evolution*, **31**, 474–483.
- 853 Kofler R, Pandey RV, Schlotterer C (2011) PoPoolation2: identifying differentiation between  
854 populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics*, **27**, 3435–  
855 3436.
- 856 Kouzarides T (2007) Chromatin modifications and their function. *Cell*, **128**, 693–705.
- 857 Kovalchuk I (2011) Transgenerational epigenetic inheritance in animals. *Frontiers in genetics*, **3**,  
858 76–76.
- 859 Kronholm I, Collins S (2015) Epigenetic mutations can both help and hinder adaptive evolution.  
860 *Molecular Ecology*, n/a–n/a.
- 861 Laland K, ULLER T, Feldman M, Sterelny K, Müller GB (2014) Does evolutionary theory need a  
862 rethink? *Nature*.
- 863 Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods*, **9**,  
864 357–359.
- 865 Lepasant JMJ, Boissier J, Climent D, Cosseau C, Grunau C (2013) Female biased sex-ratio in  
866 *Schistosoma mansoni* after exposure to an allopatric intermediate host strain of *Biomphalaria*  
867 *glabrata*. *Experimental parasitology*, **135**, 350–356.
- 868 Lepasant JMJ, Roquis D, Emans R *et al.* (2012) Combination of de novo assembly of massive  
869 sequencing reads with classical repeat prediction improves identification of repetitive  
870 sequences in *Schistosoma mansoni*. *Experimental parasitology*, **130**, 470–474.
- 871 Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform.

- 872 *Bioinformatics*, **25**, 1754–1760.
- 873 Li H, Handsaker B, Wysoker A *et al.* (2009) The Sequence Alignment/Map format and SAMtools.  
874 *Bioinformatics*, **25**, 2078–2079.
- 875 Li Z, Nie F, Wang S, Li L (2011) Histone H4 Lys 20 monomethylation by histone methylase SET8  
876 mediates Wnt target gene activation
- 877 Liu T, Rechtsteiner A, Egelhofer TA *et al.* (2011) Broad chromosomal domains of histone  
878 modification patterns in *C. elegans*. *Genome Research*, **21**, 227–236.
- 879 Makino T, McLysaght A, Kawata M (2013) Genome-wide deserts for copy number variation in  
880 vertebrates. *Nature Communications*, **4**, 1–10.
- 881 Mammana A, Chung H-R (2015) Chromatin segmentation based on a probabilistic model for read  
882 counts explains a large portion of the epigenome. *Genome Biology*, 1–12.
- 883 Martins AV (1957) Non-human vertebrate hosts of *Schistosoma haematobium* and *Schistosoma*  
884 *mansoni*. *Bulletin of The World Health Organization*, **18**, 931–944.
- 885 Mitta G, Adema CM, Gourbal B, Loker ES, Theron A (2012) Compatibility polymorphism in  
886 snail/schistosome interactions: From field to theory to molecular mechanisms. *Developmental*  
887 *& Comparative Immunology*, **37**, 1–8.
- 888 Nilsson EE, Skinner MK (2015) Environmentally induced epigenetic transgenerational inheritance  
889 of disease susceptibility. *Translational Research*, **165**, 12–17.
- 890 Oliveira G, Rodrigues NB, Romanha AJ, Bahia D (2004) Genome and genomics of schistosomes.  
891 *Canadian Journal of Zoology*, **82**, 375–390.
- 892 Panagiotou OA, Ioannidis JPA, for the Genome-Wide Significance Project (2012) What should the  
893 genome-wide significance threshold be? Empirical replication of borderline genetic  
894 associations. *International journal of epidemiology*, **41**, 273–286.
- 895 Peaston AE, Whitelaw E (2006) Epigenetics and phenotypic variation in mammals. *Mammalian*  
896 *genome : official journal of the International Mammalian Genome Society*, **17**, 365–374.
- 897 Portela J, Grunau C, Cosseau CL *et al.* (2010) Whole-genome in-silico subtractive hybridization  
898 (WISH) - using massive sequencing for the identification of unique and repetitive sex-specific  
899 sequences: the example of *Schistosoma mansoni*. *BMC Genomics*, **11**, 387.
- 900 Protasio AV, Tsai IJ, Babbage A *et al.* (2012) A systematically improved high quality genome and  
901 transcriptome of the human blood fluke *Schistosoma mansoni* (KF Hoffmann, Ed.). *PLoS*  
902 *Neglected Tropical Diseases*, **6**, e1455.
- 903 Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features.  
904 *Bioinformatics*, **26**, 841–842.
- 905 Roger E, Mitta G, Moné Y *et al.* (2008) Molecular determinants of compatibility polymorphism in  
906 the *Biomphalaria glabrata/Schistosoma mansoni* model: New candidates identified by a global  
907 comparative proteomics approach. *Molecular and Biochemical Parasitology*, **157**, 205–216.
- 908 Roquis D, Lepesant JM, Picard MAL *et al.* (2015) The epigenome of *Schistosoma mansoni* provides  
909 insight about how cercariae poise transcription until infection (SC Oliveira, Ed.). *PLoS*  
910 *Neglected Tropical Diseases*, **9**, e0003853.
- 911 Roquis D, Lepesant JM, Villafan E *et al.* (2013) Exposure to hycanthone alters chromatin structure  
912 around specific gene functions and specific repeats in *Schistosoma mansoni*. *Frontiers in*  
913 *genetics*, **5**, 207–207.

- 914 Schmitz RJ, Schultz MD, Lewsey MG *et al.* (2011) Transgenerational Epigenetic Instability Is a  
915 Source of Novel Methylation Variants. *Science (New York, N.Y.)*, **334**, 369–373.
- 916 Schotta G, Sengupta R, Kubicek S *et al.* (2008) A chromatin-wide transition to H4K20  
917 monomethylation impairs genome integrity and programmed DNA rearrangements in the  
918 mouse. *Genes & Development*, **22**, 2048–2061.
- 919 Shlyueva D, Stampfel G, Stark A (2014) Transcriptional enhancers: from properties to genome-  
920 wide predictions. *Nature Reviews Genetics*, **15**, 272–286.
- 921 Sims JK, Houston SI, Magazinnik T, Rice JC (2006) A trans-tail histone code defined by  
922 monomethylated H4 Lys-20 and H3 Lys-9 demarcates distinct regions of silent chromatin. *The*  
923 *Journal of biological chemistry*, **281**, 12760–12766.
- 924 Skinner MK, Guerrero-Bosagna C, Haque MM (2015) Environmentally induced epigenetic  
925 transgenerational inheritance of sperm epimutations promote genetic mutations. *Epigenetics*,  
926 **10**, 762–771.
- 927 Stark R, Brown G (2013) DiffBind: differential binding analysis of ChIP-Seq peak data.
- 928 Stein LD, Mungall C, Shu S, Caudy M, Mangone M, Day A, Nickerson E, Stajich JE, Harris TW, Arva,  
929 A, Lewis S (2002) The Generic Genome Browser: a building block for a model organism system  
930 database. *Genome Research*, **12**, 1599–1610.
- 931 Szklarczyk D, Franceschini A, Wyder S *et al.* (2015) STRING v10: protein-protein interaction  
932 networks, integrated over the tree of life. *Nucleic Acids Research*, **43**, D447–D452.
- 933 Tellier A, Moreno-Gómez S, Stephan W (2014) Speed of adaptation and genomic footprints of  
934 host-parasite coevolution under arms race and trench warfare dynamics. *Evolution*, **68**, 2211–  
935 2224.
- 936 Theron A, Rognon A, Gourbal B, Mitta G (2014) Multi-parasite host susceptibility and multi-host  
937 parasite infectivity: a new approach of the *Biomphalaria glabrata/Schistosoma mansoni*  
938 compatibility polymorphism. *Infection, Genetics and Evolution*, **26**, 80–88.
- 939 Thorvaldsdottir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-  
940 performance genomics data visualization and exploration. *Briefings in Bioinformatics*, **14**, 178–  
941 192.
- 942 Thrall PH, Hochberg ME, Burdon JJ, Bever JD (2007) Coevolution of symbiotic mutualists and  
943 parasites in a community context. *Trends in Ecology & Evolution*, **22**, 120–126.
- 944 Tie F, Banerjee R, Stratton CA *et al.* (2009) CBP-mediated acetylation of histone H3 lysine 27  
945 antagonizes *Drosophila* Polycomb silencing. *Development*, **136**, 3131–3141.
- 946 Trojer P, Reinberg D (2007) Facultative heterochromatin: is there a distinctive molecular  
947 signature? *Molecular Cell*, **28**, 1–13.
- 948 van der Graaf A, Wardenaar R, Neumann DA *et al.* (2015) Rate, spectrum, and evolutionary  
949 dynamics of spontaneous epimutations. *Proceedings of the National Academy of Sciences*,  
950 **112**, 6676–6681.
- 951 Vielle A, Lang J, Dong Y *et al.* (2012) H4K20me1 contributes to downregulation of X-linked genes  
952 for *C. elegans* dosage compensation (WG Kelly, Ed.). *PLoS Genetics*, **8**, e1002933.
- 953 Wan ES, Qiu W, Baccarelli A *et al.* (2012) Cigarette smoking behaviors and time since quitting are  
954 associated with differential DNA methylation across the human genome. *Human Molecular*  
955 *Genetics*, **21**, 3073–3082.



- 956 Wang Z, Zang C, Rosenfeld JA *et al.* (2008) Combinatorial patterns of histone acetylations and  
957 methylations in the human genome. *Nature Genetics*, **40**, 897–903.
- 958 Wijayawardena BK, DeWoody JA, Minchella DJ (2015) The genomic proliferation of transposable  
959 elements in colonizing populations: *Schistosoma mansoni* in the new world. *Genetica*, 1–12.
- 960 Xie C, Tammi MT (2009) CNV-seq, a new method to detect copy number variation using high-  
961 throughput sequencing. *BMC Bioinformatics*, **10**, 80.
- 962 Zentner GE, Henikoff S (2013) Regulation of nucleosome dynamics by histone modifications.  
963 *Nature Publishing Group*, **20**, 259–266.
- 964 Zhuang J, Wang J, Theurkauf W, Weng Z (2014) TEMP: a computational method for analyzing  
965 transposable element polymorphism in populations. *Nucleic Acids Research*, **42**, 6826–6838.

966

967 **Data Accessibility**

- 968 RNA-Seq and ChIP-Seq rawdata are available as fastq files under study accession number  
969 SRP035609 (BioProject Accession number PRJNA236156) at the SRA of the NCBI.