Towards an understanding of the epigenetics of schistosomes: a comparative epigenomic study

Julie Mireille and Joé Lepesant, Christoph Grunau, Céline Cosseau/+

Laboratoire Ecologie et Evolution des Interactions, Université de Perpignan Via Domitia,
52 Avenue Paul Alduy 66860, UMR 5244 CNRS Perpignan, France

As in perhaps all eukaryotes, schistosomes use a supplementary information transmitting system, the epigenetic inheritance system, to shape genetic information and to produce different phenotypes. In contrast to other important parasites, the study of epigenetic phenomena in schistosomes is still in its infancy. Nevertheless, we are beginning to grasp what goes on behind the epigenetic scene in this parasite. We have developed techniques of native chromatin immunoprecipitation (N-ChIP) and associated the necessary bioinformatics tools that allow us to run genome-wide comparative chromatin studies on Schistosoma mansoni *at different stages of its life cycle, on different strains and on different sexes. We present here an application of such an approach to study the genetic and epigenetic basis for a phenotypic trait, the compatibility of* S. mansoni *with its invertebrate host* Biomphalaria glabrata. *We have applied the ChIP procedure to two strains that are either compatible or incompatible with their intermediate host. The precipitated DNA was sequenced and aligned to a reference genome and this information was used to determine regions in which both strands differ in their genomic sequence and/or chromatin structure. This procedure allowed us to identify candidate genes that display either genetic or epigenetic difference between the two strains.*

In eukaryotes, genotypes and epigenotypes are manifested in a complex nuclear structure called chromatin. This nucleoprotein structure contains histone and non-histone proteins that interact with the genomic DNA. Chromatin exists either as a relaxed structure that is permissive to gene expression, called euchromatin, or as a condensed structure that is typically silent, called heterochromatin (Luijsterburg et al. 2008). The genotype refers to the DNA sequence of an organism. "Epigenotype" refers to the sum of chemical modifications of DNA (e.g., DNA methylation), DNA-associated proteins (e.g., histones) that can affect chromatin structure and other mediators of epigenetic information such as non-coding RNA and gene location within the nucleus (Dillon 2008, Umlauf et al. 2008). Posttranslational modifications of histones (e.g., methylation, acetylation and phosphorylation) are widely studied and have been shown to play a role in chromatin compaction and control of gene transcription (Lee et al. 2010). Orthologues of all histones apart from MacroH2A are predicted in the current genome assembly of *Schistosoma mansoni* and our laboratory and others have confirmed the presence of histone H3 and H4 and the isoforms H3K4Me3, H3S10P (unpublished observations), H3K4Me2, H3K9Ac, H3K9Me3, H3K27Me1 and H4Ac (Souza et al. 2002, Caby & Pierce 2009, Cosseau

et al. 2009). An analysis of data available in SchistoDB (Zerlotini et al. 2009) indicates that genes for all canonical histone modifying enzymes can be found.

A major advancement in the study of histone modifications in *Schistosoma* is the chromatin immunoprecipitation (ChIP) assay. ChIP allows the analysis of protein-DNA interaction in vivo. The N-terminal ends of *Schistosoma* histones H3 and H4 are conserved with other model species and, therefore, commercial antibodies can be used to immunoprecipitate modified histones and identify DNA to which they are bound. Cross-linked and native ChIP (N-ChIP) have been successfully applied to schistosomes (Caby & Pierce 2009, Cosseau et al. 2009). Massive sequencing analysis of the immunoprecipitated DNA (ChIP-Seq) and alignment against the *S. mansoni* genome has afforded a whole genome approach to visualise enrichment of specific epigenetic marks at specific *Schistosoma* loci. ChIP-Seq analysis can be performed using different antibodies that recognise different epigenetic marks responsible for euchromatic or heterochromatic structures. This technique allows the comparison of chromatin structure between life cycle stages (Cosseau et al. 2009), male or female adults (unpublished observations) or different strains. This comparative epigenomic approach has highlighted differences in chromatin structure that may impact the physiology of *S. mansoni*, such as development, sexual dimorphism and adaptation to environmental change.

In this study we have used such a comparative approach on two strains of parasites that are either compatible (C) or incompatible (IC) with a mollusc originally isolated in Brazil.

Compatibility polymorphism is a phenomenon that is observed in the interaction between *Biomphalaria glabrata* and *S. mansoni*, where some snail/schistosome

combinations are compatible and others are not (Theron & Coustau 2005). In the C interaction, the parasite manages to infect the host whereas in the IC interaction, the majority of parasites fail to infect the host. It has been hypothesised that the success or the failure of *B. glabrata*/*S. mansoni* infection relies on the matched or mismatched status of the host and parasite phenotypes. Genetic polymorphism and/or epigenetic distinction between the two strains may be responsible for this matched or mismatched status. Some molecular determinants of the compatibility polymorphism have already been described (Roger et al. 2008a, c).

In this study, we have mapped histone H3 acetylation on lysine 9 (H3K9Ac), which is a mark generally detected in active promoters (Lee et al. 2010), in both the C and IC strains of miracidia. We used an *in silico* comparison approach to identify genomic regions in which H3K9Ac enrichment differs between both strains. This method, termed ChIPDIFF, combined with quantitative polymerase chain reaction (qPCR) allowed us to highlight loci that display either genetic polymorphism or differential acetylation between the two strains.

MATERIALS AND METHODS

Parasite strains and cell culture - A C strain (Brazilian strain) and an IC strain (Guadelupian strain) toward the same mollusc (*B. glabrata* isolated from Brazil) were used in this study. Each strain was maintained in its sympatric *B. glabrata* strain and in hamsters (*Mesocricetus auratus*) as described previously (Dissous et al. 1981, Theron et al. 1997). Eggs were axenically recovered from 50-day infected hamster livers and miracidia were hatched from eggs as previously described (Theron et al. 1997). Miracidia were concentrated by sedimentation on ice for 15 min and stored at -80ºC until ChIP experiments.

Chromatin status analysis by ChIP - N-ChIP was performed as previously described (Cosseau et al. 2009) (methdb.univ-perp.fr/cgrunau/methods/native_chip_sm.html). Briefly, chromatin was extracted from both C and IC miracidia and fragmented with micrococcal nuclease. An antibody against modified histones was used to precipitate the fragmented chromatin (Table I). Immunoprecipitated DNA was purified using phenol-chloroform concentrated by precipitation and finally suspended in 80 µL Tris-ethylenediamine tetraacetic acid 10 mM.

Generation of Illumina/Solexa Libraries/processing 1G data - Generation of Illumina/Solexa Libraries was performed as previously described (Cosseau et al. 2009). The immunoprecipitated DNA was sequenced at the Oregon State University Centre for Genome Research and Biocomputing core facility using an Illumina 1G analyzer. Sequenced DNA was mapped to the *S. mansoni* reference genome (ftp://ftp.sanger.ac.uk/pub/pathogens/Schistosoma/mansoni/genome/Assembly-v3.1/) with SOAP2 software (Li et al. 2009) to generate a list of reads and their genomic locations. FindPeaks (Fejes et al. 2008) and CASHX (Fahlgren et al. 2009) were used for peak calling and generation of histograms in a genome browser (GBrowse) (Stein et al. 2002) as previously described (Cosseau et al. 2009).

In silico comparative analysis - To identify regions of statistically significant differences in H3K9Ac occurrence, we used the ChIPDiff software (Xu et al. 2008). The SOAP2 output was reformatted into the ChIPDiff input format using a custom Perl script. ChIPDiff parameters were as follows: maxIterationNum 500, minP 0.60, maxTrainingSeqNum 10000, minFoldChange 1.5, minRegionDist 500. The resulting output file was used to guide visual inspection on GBrowse. Regions of interest were selected if putative genes were in proximity.

qPCR analysis - The primers for qPCR were designed with the primer3 plus web-based interface (bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi). The genomic sequences from the reference genome (ftp://ftp.sanger.ac.uk/pub/pathogens/Schistosoma/mansoni/genome/Assembly-v3.1/) were used as a template. Primer sequences, scaffold position, expected product size and efficiency are listed in Table II. qPCR amplification was performed using 2.5 µL immunoprecipitated DNA in a final volume of 10 µL on a LightCycler® 480 Real-Time Instrument (3 µL H$_2$0, 0.5 µM of each primer, 5 µL of master mix). The following protocol was used: denaturation, 95ºC 10 min, amplification and quantification (repeat 40 times), 95ºC for 10 s, 60ºC for 10 s and 72ºC for 20 s, melting curve, 65-97ºC with a heating rate of 0.11ºC/s and continuous fluorescence measurement. For each reaction, the crossing point cycle threshold (Ct) was determined using the "second derivative" method of the LightCycler® 480 Software release 1.5. Reactions were performed in duplicate and the mean Ct was calculated. DNA recovered in the immunoprecipitated fraction was quantified by calculat-

TABLE I

Antibodies used for native chromatin immunoprecipitation (N-ChIP)

Antibody	Host	Product	Lot	Saturating quantity used for N-ChIP (µL)
H3K9Ac	Rabbit	Upstate, 07352	DAM1394804 for ChIP-seq DAM1576933 for ChIP-qPCR	8
H3K9Met3	Rabbit	Abcam, Ab8898	733951	4
H3K4Met3	Rabbit	Upstate, 04745	JBC1355340	4

qPCR: quantitative polymerase chain reaction.

TABLE II

Primers used for quantitative polymerase chain reaction (qPCR) in this study

Target	Forward (5'-3')	Reverse (5'-3')	Efficiency (IC strain)	Efficiency (C strain)	Amplicon length (bp)
Smp_173370	AGGGGGTCCTCCCTCGAAACA Smp_scaff000539:15007	CGCGCAAGTAGGCAGATAGCCA Smp_scaff000539:15143	1.95	ND	136
Smp_171100	CCTGGAATGGAGAGCTCGTATGACA Smp_scaff000421:12602	GGAGTGTGGATTGAACGCGTCG Smp_scaff000421:12724	1.91	ND	122
Smp_159800	CCGCCTGTGAACACCACTCAGC Smp_scaff000234:2330	TGCTGGACATTTTAGTTACAGTTCGCA Smp_scaff000234:2352	2	ND	122
Smp_159810	GCTTAAGTGACACTGCACAGCGA Smp_scaff000234:63585	TGGTACTCTGGATAAGCTACCGTGC Smp_scaff000234:63731	1.94	ND	146
Smp_159820	AGGCAGCCAGTCACTTGAAAGC Smp_scaff000234:74614	CGAACAGGTTCCGATGAGGACGA Smp_scaff000234:74748	2	ND	134
Smp_159830	AGCAATTGTTGTGCACCATCGACA Smp_scaff000234:161592	TCAACTCTCCCAAGGCCGTTCA Smp_scaff000234:161723	1.92	ND	131
Smp_159840	AGCTTCGACGGCTTGTTGCG Smp_scaff000234:217554	GCCCATGCACCAGAAGAGGCA Smp_scaff000234:217688	1.92	ND	134
Smp_scaff000131	TGCGCTTCCTCATCGGCCTT Smp_scaff000131:58261	TCGTGTTCATAGGTGTGTAGTGTGGT Smp_scaff000131:58367	2	ND	106
Smp_scaff000154.2	CCGGGCCGTGACCGACCTAT Smp_scaff000154:384067	TGCACTGCGAAAGGCTCGACA Smp_scaff000154: 384167	2	2	100
Smp_scaff000154.3	TTCCACGGGATGGGGTCGCT Smp_scaff000154:532853	GGGCCCCTAATTAACTTCGCCTGT Smp_scaff000154:532960	1.94	1.95	107
Smp_scaff000154.5	ACGGTGGCTTGAGGGATCAAGAA Smp_scaff000154:519935	ACAGTACGGCGGTTCCTGCG Smp_scaff000154:519465	ND	1.96	130

the region targeted is indicated by a gene name or a scaffold name when no hypothetical open reading frame has been annotated. Primer sequence is given from 5'-3' DNA strand. The 5'position on the scaffold is indicated. The qPCR efficiency has been determined when possible on the incompatible (IC) and compatible (C) strains. ND: not determined.

ing the percent input recovery (%IR) normalised with the %IR obtained at a reference locus (αTub or 28S) as previously described (Cosseau et al. 2009). Correct melting curves were checked using the Tm calling method of the LightCycler® 480 Software release 1.5. The amplification of a unique band was verified by 2% agarose gel electrophoresis for each qPCR product.

RESULTS

Identification of candidate genetic or epigenetic differences between isolates by ChIP-Seq - Our N-ChIP procedure was applied to miracidia from the C and IC strains. We then performed genome-wide sequencing of the immunoprecipitated DNA using an Illumina/Solexa 1G Genome Analyzer and aligned it to the reference genome. There were matches in 3,131,677 and 3,367,822 reads for the C and IC strain, respectively, corresponding to roughly 120 Mb.

ChIPDiff was used to identify differences in aligned sequence between the two strains. Our parameters instructed the software to screen the immunoprecipitated aligned sequences in 1 kb blocks and highlight those displaying more than a 1.5 fold change in H3K9 acetylation. The C, IC and reference strains are three different *S. mansoni* isolates. If the genetic polymorphism between these strains is greater than the mismatches allowed for alignment (> 2 in 36 bp), then a difference in sequence alignment may occur at some loci. In this case, the ChIPDiff method allows for identification of genetic polymorphism between the C and IC strains. If there are differences in copy number between these strains (duplicated sequences), then an apparent difference in enrichment of sequences may be observed. If there are no genetic differences and no copy number effects, then the differences detected by ChIPDiff are due to differences in epigenetic marks. It should be pointed out that epigenetic differences may accumulate with genetic differences. Conversely, an epigenetic difference may be compensated by a genetic difference. In these cases, the ChIPDiff method will not be able to distinguish between these scenarios.

Using the ChIPDiff method, we identified 321 differences between the two strains. There were 128 "1 kb blocks" that were enriched for sequences from the IC strain and the remaining 193 "1 kb blocks" were enriched for sequences from the C strain (Fig. 1). We discarded the individual "1 kb blocks" and only further considered regions that contained clusters of "1 kb blocks" spanning several kb regions. We selected nine regions, some of which contain hypothetical annotated genes. Results are listed in Table III. Of particular interest, we identified a 300 kb region on Smp_scaff000234 with more aligned sequences in the IC strains compared with the C strains. This region contains four genes encoding for Egg Secreted Proteins (ESP15 family) recently classified as MEG2 family genes [micro exon genes (MEG)] (DeMarco et al. 2010). Another region enriched in IC sequences contains a gene highly expressed in eggs and miracidia that encodes a metalloprotease from the neprelysin-2 family. Most of the region with more aligned sequences in the C strains contains genes that encode for hypothetical proteins.

Confirmation of genetic and epigenetic difference between the C and IC strains by qPCR - To determine whether the observed difference in aligned sequences was due to epigenetic or genetic reasons, we further performed qPCR experiments. We selected 11 loci for which ChIPDiff had identified a difference in aligned sequences between the C and IC strains. These loci may or may not be located in a predicted open reading frame. Primers were designed to target these loci and the %IR was calculated for each position. Results are presented in Table IV. Six primer sets amplified the ChIP products from the IC strain with good efficiency, but did not amplify anything from the C strains. Conversely, one primer set amplified the ChIP products from the C strain, but no amplification could be obtained from the IC strain. For these positions, we hypothesised that the targets are located in regions where genetic polymorphism between the two strains occurs, which is the case for the region containing the MEG2 family genes. For two regions, qPCR did not allow for confirmation of an observed difference by ChIPDiff with no apparent genetic polymorphism. However, a copy number effect was detected for these two regions. Finally, we confirmed a difference in H3K9Ac enrichment in the IC strains compared with the C strain in the Smp_171100 and Smp_scaff000154.3 regions, which indicates that these two loci are more euchromatic in the IC strains compared with the C strain. For the Smp_171100 locus, a heterochromatic epigenetic mark [H3 trimethylated on lysine 9 (H3K9Me3)] was further tested by ChIP-qPCR and it was confirmed that this region is more euchromatic in the IC strains compared with the C strain (Fig. 2).

DISCUSSION

Throughout its life cycle, *S. mansoni* encounters and must adapt to different types of environments. Interaction with host partners is an example of adaptation to such a changing environment. The compatibility polymorphism described for the *S. mansoni*/*B. glabrata* interaction is an illustration of such an adaptation displayed by the
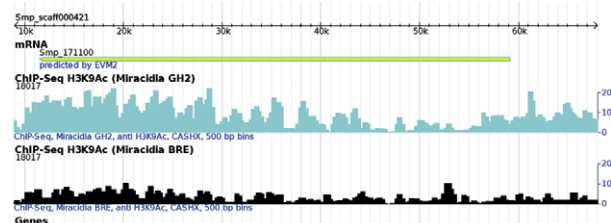

Fig. 1: Smp_scaff000421:9000..68000. Immunoprecipitated sequences using an antibody that recognizes H3 acetylated on lysine 9 (H3K9Ac) was aligned against *Schistosoma mansoni* reference genome. Aligned sequences can be visualized using an inhouse *S. mansoni* genome browser. The figure represents the data from Smp_scaff000421, position 9000..68000. mRNA represents a predicted gene (Smp_171100) in this region. Chromatin immunoprecipitation (ChIP)-Seq H3K9Ac (miracidiae GH2 and BRE) represents the number of hits in a 500 bp (CASHX data) window obtained for the incompatible (IC) and compatible (C) strains after ChIP using an antibody against H3 acetylated on lysine 9.

parasite toward its invertebrate host (Theron & Coustau 2005). Molecular candidates for this compatibility polymorphism have been previously described (Roger et al. 2008a, c). In this work, we further raised the question of the implications of genetic vs. epigenetic information for the compatibility polymorphism phenotype. We immunoprecipitated chromatin from two strains, C and IC, being compatible and incompatible, respectively, with the mollusk *B. glabrata*. We used an antibody against histone H3 acetylated on lysine 9, which is a mark found in euchromatin, and we compared the ChIP-Seq profile of these two strains by ChIPDiff. This method allowed us to identify at least nine regions that displayed quantitative differences in aligned sequences against the reference genome, thus highlighting either genetic or epigenetic differences between the two strains. To confirm the genetic and/or epigenetic differences, we used a qPCR approach. We assumed that no PCR amplification in one strain (with a positive result in the other strain) means that there is genetic polymorphism between these two strains. We identified seven loci with a genetic difference, two loci with a difference in copy number and two other loci with an epigenetic difference.

Among the genetic candidates, we identified a 300 kb region, located on Smp_scaff000234, that contains four copies of genes annotated as MEG. These MEG-type genes have only been described in *S. mansoni* (DeMarco et al. 2010). They are composed of 18 different family members and are a molecular system for creating protein variation through alternate splicing of short exons. The MEG-2 member is detected in our study as a genetic candidate for the compatibility polymorphism phenotype. It is expressed in eggs and miracidia and has been shown to be glycosylated and secreted in eggs. Its potential role in the context of immune evasion strategy has been previously discussed (DeMarco et al. 2010). It is of particular interest that MEG-2 shares many common features with the *SmPoMuc* genes, previously described as major molecular determinants for the compatibility polymorphism (Roger et al. 2008b). *SmPoMuc* genes encode highly polymorphic glycosylated proteins that are secreted in miracidia and their cDNA variants seem to be generated by the "pick and mix" strategy as described for the MEG genes (Berriman et al. 2009). *SmPoMuc* genes display MEG-like characteristics and their potential role as polymorphic antigenic variants in

TABLE III

Regions containing a difference in the number of aligned sequences between the compatible (C) and incompatible (IC) strains

Scaffold	Positions	Predicted open reading frame	Annotation	Expression
		IC > C		
Smp_scaff000234	0-310 kb	Smp_159800	MEG2 family (ESP15)	Egg
		Smp_159810	MEG2 family (ESP15)	Egg
		Smp_159820	MEG2 family (ESP15)	Egg
		Smp_159830	MEG2 family (ESP15)	Egg
		Smp_159840	Putative dynein heavy chain	Egg, schistosomula, adults
Smp_scaff000421	10-70 kb	Smp_171100	Zinc metalloprotease	Egg, miracidia, female
Smp_scaff000496	0-2.3 kb and 10-15 kb	None		
		C > IC		
Smp_scaff000131	0-70 kb	Smp_149750	Hypothetical protein	Egg, germball, cercaria
		Smp_051690	Hypothetical protein	No EST in database
		Smp_149760	Hypothetical protein	No EST in database
Smp_scaff000490	8-16 kb	None		
Smp_scaff000538	0-24 kb	Smp_093470	Hypothetical protein	No EST in database
Smp_scaff000539	0-35 kb	Smp_173370	M13 unassigned peptidase	No EST in database
Smp_scaff019017	10-40 kb	Smp_121530	Hypothetical protein	Germball, adult
		Alternative IC > C/C > IC		
Smp_scaff000154	449-500 kb	Smp_152600	Paired box protein pax-6	Egg, miracidium
		Smp_152610	Hypothetical protein	No EST in database
		Smp_152620	Hook homolog	Egg, germball, schistosomula
		Smp_152630	Hypothetical protein	No EST in database

nine regions that display a difference between the C and IC strains in aligned sequences have been selected. Positions of the concerned region on the selected scaffold are indicated. Information on putative genes contained in these regions, their annotations and their corresponding expressed sequence tag (EST) are given (available from the SchistoDB: schistodb.net/schistodb20/). MEG: micro exon genes.

the context of an arm race hypothesis has been discussed (Moné et al. 2010). These features make the MEG-2 genes promising candidates for a second determinant of the compatibility polymorphism phenotype.

Among the candidates with an epigenetic difference, we identified a 60 kb region on Smp_scaff000421 that contains a gene (Smp_171100) annotated as a metalloprotease of the neutral endopeptidase (NEP) family. The role of *S. mansoni* NEP for the immuno-modulation of *B. glabrata* has been previously described (Duvaux-Miret et al. 1992). This gene is expressed in eggs and miracidia and we showed in this study that it is located in a region that is more euchromatic in the IC strain compared with the C strain. Therefore, this gene constitutes an excellent candidate to study the epigenetic bases of the compatibility polymorphism phenotype.

While the role of epigenetics for adaptation, development and immune evasion strategy has been largely described in other parasite models (Dixon et al. 2010, Merrick & Duraisingh 2010, Wang et al. 2010), epigenetic studies in *S. mansoni* are still in their infancy. Development of tools to study the role of histone modifications and the recent release of the *S. mansoni* genome (Berriman et al. 2009) have allowed for performing genome-wide epigenomic studies on this organism. In this study, we applied such tools to study adaptation to invertebrate hosts. This relatively straightforward approach that consists of a combination of low-coverage ChIP-Seq, bioinformatics analysis and confirmation by qPCR has

enabled highlighting of candidates for the compatibility polymorphism phenotypes that we believe are relevant for immune evasion strategies.

Perspectives - In addition to the study of the histone modifications presented in this paper, it will be of particular interest to broaden the study of schistosome's chromatin structure to other epigenetic events. Of particular interest is the study of non-coding RNA. We recently identified female specific satellite DNA se-
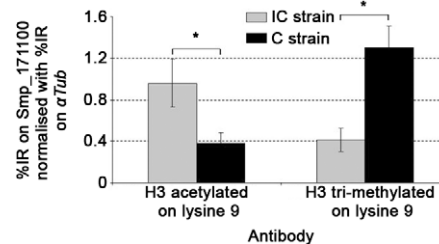

Fig. 2: percent input recovery (%IR) of gDNA in the compatible (C) and incompatible (IC) strains after chromatin immunoprecipitation (ChIP). ChIP was applied on C and IC miracidiae gDNA using two antibodies that recognize H3 acetylated on lysine 9 and H3 tri-methylated on lysine 9. Quantitative polymerase chain reaction was used to quantify the %IR on the target locus (Smp_171100) and reference genes (αTub) as indicated in the method section. Results represent the mean of three biological repeats of the %IR of Smp_171100 normalised with αTub. Asterisk means a t-test result below 0.05 between BRE and GH2 data.

TABLE IV

Comparison of chromatin immunoprecipitation-quantitative polymerase chain reaction (ChIP-qPCR) and sequencing analysis of the immunoprecipitated DNA (ChIP-Seq) results

Target	Number of loci	PCR on IC	PCR on C	Differences between C and IC strains using ChIP-qPCR	Differences between C and IC strains using ChIP-Seq	Comparison of ΔCt(IC) and ΔCt(C)	Genetic or epigenetic?
Smp173370	2	+	+	IC > C	C > IC	ΔCt(IC) > ΔCt(C)	No confirmation
Smp171100	2	+	+	IC > C	IC > C	ΔCt(IC) = ΔCt(C)	Epigenetic
Smp159800	2	+	-	ND	IC > C	ND	Genetic
Smp159810	4	+	-	ND	IC > C	ND	Genetic
Smp159820	2	+	-	ND	IC > C	ND	Genetic
Smp159830	1	+	-	ND	IC > C	ND	Genetic
Smp159840	1	+	-	ND	IC > C	ND	Genetic
Scaff000131	1	+	+	IC = C	C > IC	ΔCt(IC) > ΔCt(C)	No confirmation
Scaff000154.2	1	+	-	ND	IC > C	ND	Genetic
Scaff000154.3	1	+	+	IC > C	IC > C	ΔCt(IC) = ΔCt(C)	Epigenetic
Scaff000154.5	1	-	+	ND	C > IC	ND	Genetic

the region targeted is indicated by a gene name or a scaffold name when no hypothetical open reading frame has been annotated. The primer sequences for qPCR amplification and accurate positions for these targets are given in Table II. The number of loci targeted by the primer sets is given in the column "loci". PCR amplification of the ChIP products is indicated for compatible (C) and incompatible (IC) strains. We assume that no amplification in one strain (but working on the other strain) means that genetic polymorphism occurs between these two strains and prevent the primer set to work correctly. Differences of immunoprecipitated DNA between the C and IC strains using H3 acetylated on lysine 9 are indicated for a detection using qPCR or Solexa sequencing. The ΔCt represents the Cycle threshold (Ct) (target) - Ct (αTub), ΔCt in the IC and C strains are compared in order to determine if the target copy number in the C and IC strains are identical (we assume that αTub copy number is the same in the C and IC strains). ND: not determined.

quences for which we have hypothesised a role in the sexual dimorphism observed at the adult stages (unpublished observations). MicroRNAs candidates have also been reported in the genome of *S. mansoni* (Simões et al. 2011) and in *Schistosoma japonicum* (Huang et al. 2009). The identification of regulatory non-coding RNA molecules in *Schistosoma* opens new avenues that will likely broaden the epigenetic understandings of schistosomes. Although DNA methylation has not been detected in schistosomes (Fantappié et al. 2001), a bioinformatic survey of the sequenced *S. mansoni* genome revealed the presence of a DNA methyl transferase (DNMT2). We anticipate that low levels of DNA methylation may occur in schistosomes that likely mediate fine-tuning of the development of this parasite at specific stages of its life cycle. This was previously described for other "DN-MT2-only" organisms (Phalke et al. 2009). It will certainly be necessary to increase the study of these other epigenetic events to decipher the impact of epigenetics for the schistosomes physiology.

REFERENCES

Berriman M, Haas BJ, LoVerde PT, Wilson RA, Dillon GP, Cerqueira GC, Mashiyama ST, Al-Lazikani B, Andrade LF, Ashton PD, Aslett MA, Bartholomeu DC, Blandin G, Caffrey CR, Coghlan A, Coulson R, Day TA, Delcher A, DeMarco R, Djikeng A, Eyre T, Gamble JA, Ghedin E, Gu Y, Hertz-Fowler C, Hirai H, Hirai Y, Houston R, Ivens A, Johnston DA, Lacerda D, Macedo CD, McVeigh P, Ning Z, Oliveira G, Overington JP, Parkhill J, Pertea M, Pierce RJ, Protasio AV, Quail MA, Rajandream MA, Rogers J, Sajid M, Salzberg SL, Stanke M, Tivey AR, White O, Williams DL, Wortman J, Wu W, Zamanian M, Zerlotini A, Fraser-Liggett CM, Barrell BG, El-Sayed NM 2009. The genome of the blood fluke *Schistosoma mansoni*. *Nature 460*: 352-358.

Caby S, Pierce RJ 2009. Quantitative chromatin immunoprecipitation (Q-ChIP) applied to *Schistosoma mansoni*. *Mol Biochem Parasitol 166*: 77-80.

Cosseau C, Azzi A, Smith K, Freitag M, Mitta G, Grunau C 2009. Native chromatin immunoprecipitation (N-ChIP) and ChIP-Seq of *Schistosoma mansoni*: critical experimental parameters. *Mol Biochem Parasitol 166*: 70-76.

DeMarco R, Mathieson W, Manuel SJ, Dillon GP, Curwen RS, Ashton PD, Ivens AC, Berriman M, Verjovski-Almeida S, Wilson RA 2010. Protein variation in blood-dwelling schistosome worms generated by differential splicing of micro-exon gene transcripts. *Genome Res 20*: 1112-1121.

Dillon N 2008. The impact of gene location in the nucleus on transcriptional regulation. *Dev Cell 15*: 182-186.

Dissous C, Dissous C, Capron A 1981. Isolation and characterization of surface antigens from *Schistosoma mansoni* schistosomula. *Mol Biochem Parasitol 3*: 215-225.

Dixon SE, Stilger KL, Elias EV, Naguleswaran A, Sullivan WJ Jr 2010. A decade of epigenetic research in *Toxoplasma gondii*. *Mol Biochem Parasitol 173*: 1-9.

Duvaux-Miret O, Stefano GB, Smith EM, Dissous C, Capron A 1992. Immunosuppression in the definitive and intermediate hosts of the human parasite *Schistosoma mansoni* by release of immunoactive neuropeptides. *Proc Natl Acad Sci USA 89*: 778-781.

Fahlgren N, Sullivan CM, Kasschau KD, Chapman EJ, Cumbie JS, Montgomery TA, Gilbert SD, Dasenko M, Backman TW, Givan SA, Carrington JC 2009. Computational and analytical framework for small RNA profiling by high-throughput sequencing. *RNA 15*: 992-1002.

Fantappié MR, Gimba ER, Rumjanek FD 2001. Lack of DNA methylation in *Schistosoma mansoni*. *Exp Parasitol 98*: 162-166.

Fejes AP, Robertson G, Bilenky M, Varhol R, Bainbridge M, Jones SJ 2008. FindPeaks 3.1: a tool for identifying areas of enrichment from massively parallel short-read sequencing technology. *Bioinformatics 24*: 1729-1730.

Huang J, Hao P, Chen H, Hu W, Yan Q, Liu F, Han ZG 2009. Genome-wide identification of *Schistosoma japonicum* microRNAs using a deep-sequencing approach. *PLoS ONE 4*: e8206.

Lee JS, Smith E, Shilatifard A 2010. The language of histone crosstalk. *Cell 142*: 682-685.

Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K, Wang J 2009. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics 25*: 1966-1967.

Luijsterburg MS, White MF, van Driel R, Dame RT 2008. The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes. *Crit Rev Biochem Mol Biol 43*: 393-418.

Merrick CJ, Duraisingh MT 2010. Epigenetics in *Plasmodium*: what do we really know? *Eukaryot Cell 9*: 1150-1158.

Moné Y, Gourbal B, Duval D, Du Pasquier L, Kieffer-Jaquinod S, Mitta G 2010. A large repertoire of parasite epitopes matched by a large repertoire of host immune receptors in an invertebrate host/ parasite model. *PLoS Negl Trop Dis 4*: e813.

Phalke S, Nickel O, Walluscheck D, Hortig F, Onorati MC, Reuter G 2009. Retrotransposon silencing and telomere integrity in somatic cells of *Drosophila* depends on the cytosine-5 methyltransferase DNMT2. *Nat Genet 41*: 696-702.

Roger E, Gourbal B, Grunau C, Pierce RJ, Galinier R, Mitta G 2008a. Expression analysis of highly polymorphic mucin proteins (SmPoMuc) from the parasite *Schistosoma mansoni*. *Mol Biochem Parasitol 157*: 217-227.

Roger E, Grunau C, Pierce RJ, Hirai H, Gourbal B, Galinier R, Emans R, Cesari IM, Cosseau C, Mitta G 2008b. Controlled chaos of polymorphic mucins in a metazoan parasite (*Schistosoma mansoni*) interacting with its invertebrate host (*Biomphalaria glabrata*). *PLoS Negl Trop Dis 2*: e330.

Roger E, Mitta G, Moné Y, Bouchut A, Rognon A, Grunau C, Boissier J, Théron A, Gourbal BE 2008c. Molecular determinants of compatibility polymorphism in the *Biomphalaria glabrata*/*Schistosoma mansoni* model: new candidates identified by a global comparative proteomics approach. *Mol Biochem Parasitol 157*: 205-216.

Simões MC, Lee J, Djikeng A, Cerqueira GC, Zerlotini A, da Silva-Pereira RA, Dalby AR, Loverde P, El-Sayed NM, Oliveira G 2011. Identification of *Schistosoma mansoni* microRNAs. *BMC Genomics 19*: 12-47.

Souza PP, Santos DN, Pena SDJ, Franco GR 2002. Cloning and molecular characterization of the *Schistosoma mansoni* genes RbAp48 and histone H4. *Mem Inst Oswaldo Cruz 97* (Suppl. I): 77-84.

Stein LD, Mungall C, Shu S, Caudy M, Mangone M, Day A, Nickerson E, Stajich JE, Harris TW, Arva A, Lewis S 2002. The generic genome browser: a building block for a model organism system database. *Genome Res 12*: 1599-1610.

Theron A, Coustau C 2005. Are *Biomphalaria* snails resistant to *Schistosoma mansoni*? *J Helminthol 79*: 187-191.

Theron A, Pages JR, Rognon A 1997. *Schistosoma mansoni*: distribution patterns of miracidia among *Biomphalaria glabrata* snail as related to host susceptibility and sporocyst regulatory processes. *Exp Parasitol 85*: 1-9.

Umlauf D, Fraser P, Nagano T 2008. The role of long non-coding RNAs in chromatin structure and gene regulation: variations on a theme. *Biol Chem 389*: 323-331.

Wang QP, Kawahara T, Horn D 2010. Histone deacetylases play distinct roles in telomeric VSG expression site silencing in African trypanosomes. *Mol Microbiol 77*: 1237-1245.

Xu H, Wei CL, Lin F, Sung WK 2008. An HMM approach to genome-wide identification of differential histone modification sites from ChIP-seq data. *Bioinformatics 24*: 2344-2349.

Zerlotini A, Heiges M, Wang H, Moraes RL, Dominitini AJ, Ruiz JC, Kissinger JC, Oliveira G 2009. SchistoDB: a *Schistosoma mansoni* genome resource. *Nucleic Acids Res 37*: D579-782.