Evolution of gene dosage on the Z-chromosome of schistosome parasites

Marion Picard, Celine Cosseau, Sabrina Ferré, Quack Thomas, C. Grevelding, Yohann Coute, Beatriz Vicoso

Full title: Evolution of gene dosage on the Z-chromosome of schistosome parasites

Short title: Dosage compensation evolution in schistosomes.

Keywords: Sex chromosome evolution, dosage compensation, Ohno's hypothesis, female heterogamety, schistosomes

Authors: Marion A. L. Picard[1], Céline Cosseau[2], Sabrina Ferré[3], Quack Thomas[4], Christoph G. Grevelding[4], Yohann Couté[3], Beatriz Vicoso[1,o]

Author information:

[1] Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg, Austria

[2] Université de Perpignan Via Domitia, IHPE UMR 5244, CNRS, IFREMER, Université de Montpellier, F-66860 Perpignan, France.

[3] Univ. Grenoble Alpes, CEA, Inserm, BIG-BGE, 38000 Grenoble, France

[4] Institute for Parasitology, BFS, Justus-Liebig-University, 35392 Giessen, Germany.

°Corresponding author: beatriz.vicoso@ist.ac.at

Abstract:

XY systems usually show chromosome-wide compensation of X-linked genes, while in many ZW systems, compensation is restricted to a minority of dosage sensitive genes. Why such differences arose is still unclear. Here, we combine comparative genomics, transcriptomics and proteomics to obtain a complete overview of the evolution of gene dosage on the Z-chromosome of *Schistosoma* parasites. We compare the Z-chromosome gene content of African (*Schistosoma mansoni* and *S. haematobium*) and Asian (*S. japonicum*) schistosomes, and describe lineage-specific evolutionary strata. We use these to assess gene expression evolution following sex-linkage. The resulting patterns suggest a reduction in expression of Z-linked genes in females, combined with up-regulation of the Z in both sexes, in line with the first step of Ohno's classic model of dosage compensation evolution. Quantitative proteomics suggest that post-transcriptional mechanisms do not play a major role in balancing the expression of Z-linked genes.

Introduction

In species with separate sexes, genetic sex determination is often present in the form of differentiated sex chromosomes (1). A sex-specific chromosome can be carried by the male (such as the Y of mammals and fruit flies, in male heterogamety) or by the female (such as the W of birds, in female heterogamety). These sex chromosomes originally arise from pairs of autosomes, which stop recombining after they acquire a sex-determining region (2,3). The loss of recombination between X/Z and Y/W chromosomes is likely driven by selective pressures to link the sex-determining gene and alleles with sexually antagonistic effects, and often occurs through inversions on the sex-specific chromosome (4,5). The inverted Y/W-linked region stops recombining entirely, which hampers the efficacy of selection and leads to its genetic degeneration (3,6). The appearance of further sexually antagonistic mutations can restart the process and select for new non-recombining regions, creating sex chromosome "strata" of different ages (7–10). Eventually, this suppression of recombination can extend to most of the chromosome, leading to gene-poor, mostly heterochromatic sex chromosomes such as the Y chromosome of mammals (11).

The loss of one gene copy on the Y/W is predicted to result in a two-fold reduction of expression in the heterogametic sex, as gene expression is correlated to gene copy number (12). This can cause imbalances in gene networks composed of both X/Z-linked and autosomal genes (13). Such imbalances can drive the appearance of dosage compensation mechanisms, which target X/Z chromosomes and regulate their expression to restore optimal dosage (14,15). While X/Z upregulation in the heterogametic sex is required to re-establish balanced levels of expression, global downregulation in the homogametic sex is also observed (e.g. X-inactivation in mammals). Ohno suggested a two-step mechanism, in which the initial upregulation of expression is not sex-specific. This leads to an excess of dosage in the homogametic sex, and secondarily selects for further repressing mechanisms (“Ohno’s hypothesis” of dosage compensation, (14)). How relevant this model is to the evolution of mammalian dosage

compensation is still under debate (e.g. 16,18,19,21,22). Independent of the underlying mechanisms, balanced gene expression between males and females in species with differentiated sex chromosomes was used as diagnostic of a chromosome-wide (also referred to as “global”, or “complete”) mechanism of dosage compensation in many different clades (21–23).

In ZW systems, the loss of genes on the sex-specific W chromosome is generally accompanied by unequal expression levels of the Z-chromosome between ZZ males and ZW females, as well as reduced expression of the Z relative to the autosomes in females. This has generally been interpreted as a lack of chromosome-wide dosage compensation (also referred to as “partial”, or “incomplete”), with individual dosage-sensitive genes being independently regulated instead. Incomplete dosage compensation was described in a wide range of species, including birds (24–27), fishes (28) and snakes (9). So far, Lepidoptera are the only exception to this observation (23,29). Why many ZW systems should fail to acquire a global mechanism of dosage compensation is not entirely clear, though several and non-mutually exclusive hypotheses have been put forward (see Discussion, and Gu and Walters (16) for a review). Another possibility is that the male-bias of the Z is instead caused by an accumulation of genes with male functions due to the male-biased transmission of the Z, which may favor the fixation of sexually antagonistic male-beneficial mutations on this chromosome.

While the direct comparison of male and female expression of X/Z-linked and autosomal genes has provided an overview of dosage compensation in many clades, it suffers from several drawbacks (23). First, chromosome-wide dosage compensation can lead to strongly sex-biased expression, if only the initial up-regulation of expression in both sexes has occurred (but not the secondary down-regulation of the Ohno hypothesis (14)). This has been suggested for the flour beetle (30) and for the young sex chromosomes of the threespine sticklebacks (31). Biases in expression levels between sexes and/or chromosomes may also have been present ancestrally, before the present sex-chromosomes evolved, and using a proxy for ancestral expression can

yield insights into the direct consequences of sex-linkage (17,32,33). Finally, the vast majority of studies relied only on microarray or RNA-seq data and did not consider any post-transcriptional regulation that might affect gene dosage at the protein level, but not at the transcript level (whereas protein dosage is in most cases the functionally relevant measure). For instance, a proteomic analysis in birds found that several genes appeared to be partially equalized at the protein level despite being strongly male-biased at the transcript level (34). In humans, post-transcriptional regulation does not appear to play a major role in dosage compensation (35).

Here we combine comparative genomics, transcriptomics and quantitative proteomics to obtain a complete overview of the evolution of gene dosage on the Z-chromosome of parasites of the genus *Schistosoma*. Schistosomes are a group of blood parasites that can cause schistosomiasis in humans (36). Their complex life cycle is characterized by a phase of clonal multiplication in an intermediate mollusk host, and a phase of sexual reproduction in the final warm-blooded host. Unlike the other 20,000 species of hermaphroditic platyhelminths, schistosomes have separate sexes: sexual reproduction occurs immediately after the male and female primary development in their definitive host, and mating is compulsory for the sexual maturation of females (37,38). Sex determination is genetic, and relies on a pair of cytogenetically well-differentiated ZW chromosomes (39). All schistosomes are thought to share the same ancestral pair of ZW sex chromosomes, but differences in their morphology and in the extent of heterochromatization of the W suggest that different strata were acquired independently by different lineages (39,40).

The model blood fluke *Schistosoma mansoni* was one of the first ZW clades to be evaluated for the presence of global dosage compensation, through the comparison of male and female microarray data derived from several tissues (41). It showed reduced expression of Z-linked genes in females relative (i) to the autosomes and (ii) to males, consistent with a lack of chromosome-wide dosage compensation. Interestingly, the reduction of Z-expression in females was less than two-fold, and the Z:autosome ratio of expression was slightly, but consistently,

111 greater than one in males. Our combined genomic, transcriptomic, and proteomic approaches
112 allow us to fully probe the evolution of the male-biased expression of the Z, and suggests a
113 more complex scenario than previously proposed. We discuss this in light of the different
114 hypotheses put forward to account for the evolution of gene dosage on Z chromosomes.
115
116 **RESULTS**
117 1. Genomic differentiation of ZW sex chromosomes in Asian and African lineages.
118 The difference in morphology of the ZW pair in African and Asian schistosomes suggests that
119 the two lineages may differ in their gene content (39). We compared the gene content of the Z-
120 chromosomes of three different species: *S. mansoni* and *Schistosoma haematobium,* which
121 belong to African schistosomes, and *Schistosoma japonicum,* an Asian schistosome (Figure 1).
122 We first identified syntenic blocks between the *S. mansoni* genome and the *S. haematobium*
123 and *S. japonicum* scaffolds. To this end, we mapped all *S. mansoni* protein coding sequences to
124 the genome assemblies of the two other species and selected only the hits with the highest
125 scores, yielding 9,504 *S. mansoni* / *S. haematobium* orthologs and 8,555 *S. mansoni* / *S.*
126 *japonicum* orthologs (Table 1). Scaffolds were then assigned to one of the *S. mansoni*
127 chromosomes, based on their ortholog content (Figure 1-source data 1).
128 We further performed a comparative coverage analysis to define the Z-specific regions of the
129 three species. Z-derived sequences are expected to display half the genomic coverage in ZW
130 females as in ZZ males, and as the autosomes. We thus mapped male and female genomic
131 reads (or only female reads in the case of *S. haematobium*) to the reference genome of each
132 species (42–45). Publicly available raw reads were used for *S. mansoni* and *S. haematobium*
133 (Wellcome Trust Sanger Institute Bioprojects PREJB2320 and PREJB2425), while male and
134 female *S. japonicum* were sequenced for this study. We then estimated the per base genomic
135 coverage. Median coverage values were 18.40 and 18.99 for *S. mansoni* male and female
136 libraries; 23.5 and 7.43 for *S. haematobium* female#1 and female#2 libraries; 23.77 and 20.53

for *S. japonicum* male and female libraries. Z-specific genomic regions were defined by a maximum value of the female:male ratio of coverage (*S. mansoni*: log2(female:male)=-0.4; *S. japonicum*: log2(female:male)=-0.84), or a maximum value of female coverage (*S. haematobium*: log2(female)=4.41). Details of how these cutoff values were obtained are provided in the Methods and Appendix 1. This analysis resulted in 285 newly described Z-specific genes in *S. mansoni* that were previously located on 19 unplaced scaffolds longer than 50kb, and to a refined pseudoautosomal/Z-specific structure of the published ZW linkage group (42,43) (Figure 1-source data 2). It further allowed us to define 379 Z-specific scaffolds (containing 1,409 annotated genes with orthologs in *S. mansoni*) in *S. haematobium* (Figure 1-source data 3 for exhaustive list) and 461 Z-specific scaffolds (containing 706 annotated orthologs) in *S. japonicum* (Figure 1-source data 4 for exhaustive list).

While the content of the Z was largely shared between the African *S. mansoni* and *S. haematobium* (Table1, Figure 1), large differences were found between the African and Asian lineages: only 476 Z-specific genes were shared by *S. mansoni* and *S. japonicum*, while 306 were only Z-specific in *S. mansoni* and 137 only in *S. japonicum* (Table1). Of all these Z-specific genes, 613 were already mapped to the *S. mansoni* ZW linkage group (42,43) and, when plotted along the Z-chromosome, outlined three different evolutionary strata: one shared ancestral stratum (S0: 367 genes) and two lineage-specific strata (S1mans, specific to the African schistosomes, with 180 genes; and S1jap, specific to *S. japonicum*, with 66 genes) (Table 1, Figure 1, and Figure 1-source data 1). The presence of pseudoautosomal regions throughout the S0 (Figure 1) is likely due to errors in the genome assembly. All further analyses were run using all newly identified Z-specific genes, but hold when only Z-specific genes that were previously mapped to the ZW linkage group are considered (Appendix 1).

Table 1. Number of orthologs assigned as Z-linked and autosomal in *S. mansoni*, *S. haematobium* and *S. japonicum*, based on the female:male (or female for *S. haematobium*) coverage patterns.

		SCHISTOSOMA JAPONICUM				*SCHISTOSOMA HAEMATOBIUM*			
	Categories	Z-specific	Autosomal	Ambiguous	Not covered	Z-specific	Autosomal	Ambiguous	Not covered
SCHISTOSOMA MANSONI	Z-specific	476 *(S0)*	306 *(S1mans)*	20	3	847	36	23	10
	Autosomal	137 *(S1jap)*	7,062	91	13	216	7,462	262	105
	Excluded	57	383	4	3	105	411	20	7
	Orthologs total	**8,555**				**9,504**			

2. Consistent patterns of expression in *S. mansoni* and *S. japonicum*.

In order to test for dosage compensation, the median expression of Z-specific genes in ZW females can be compared to the median autosomal expression (Z:AA ratio) and/or to the Z-specific gene expression in ZZ males (F:M ratio). In females, a Z:AA or F:M ratio of ~1 supports global dosage compensation, while a ratio between 0.5 and 1 suggests partial or local dosage compensation. We performed this analysis in *S. mansoni* and in *S. japonicum*, using publicly available RNA-seq reads derived from a sexually undifferentiated stage (schistosomula, (46,47)) and a sexually mature stage (adults, Wellcome Trust Sanger Institute Bioproject PRJEB1237, (47)). The inclusion of a sexually immature stage is important, as much of the expression obtained from adults will necessarily come from their well-developed gonads. Sex-linked genes are often sex-biased in the germline, even in organisms that have chromosome-wide dosage compensation (e.g. due to sex-chromosome inactivation during gametogenesis), and the inclusion of gonad expression has led to inconsistent assessments of the status of dosage compensation in other clades (16,29). Reads were mapped to their respective genomes, and expression values in Reads Per Kilobase Million (RPKM) were calculated for each gene (Figure

2, Figure 2-source data 1 and 2); only genes with a minimum RPKM of 1 in both sexes were considered.

We consistently observed a strong male bias in the expression of Z-specific genes in both stages and species (F:M ratio between 0.58 and 0.69, Figure 2 and Supplementary File 1), consistent with local or incomplete dosage compensation. While this was generally supported by the lower expression levels of Z-specific genes in females when compared to the autosomes (Z:AA ratio between 0.73 and 0.85; Figure 2, Supplementary File 1), this difference was only apparent for some filtering procedures (Figure 2-figure supplement 1 to Figure 2-figure supplement 14), and even then was not sufficient to fully account for the strong male-bias of the Z. Instead, the higher expression of the male Z in both stages and species (ZZ:AA ratio between 1.25 and 1.46, Supplementary File 1) appeared to also contribute to the male-bias of Z-linked genes. These patterns were qualitatively robust to changes in the methods used to estimate expression (RPKM or TPM [Transcripts Per Kilobase Million]), in the filtering procedure (RPKM>0, RPKM>1, TPM>0 or TPM>1), and when only genes that were previously mapped to the ZW linkage group were considered. These analyses were further performed independently in the S0, S1mans and S1jap strata, which showed no significant difference in the extent of their male bias. All the resulting plots are shown in Figure 2-figure supplement 1 to Figure 2-figure supplement 14. Finally, Z-specific genes were found to be male-biased even when only genes with broad expression were considered (RPKM>1 and RPKM>3 in all samples, and when genes with strong sex-biases in expression were excluded (M:F>2 or F:M>2, Figure 2-figure supplement 15 and Figure 2-figure supplement 16, I-L panels), confirming that this pattern does not appear to be driven simply by the presence of genes with sex-specific functions on the Z-chromosome. No further influence of known protein-protein interactions was detected (Figure 2-figure supplement 17, Appendix 1).

208 3. Convergent upregulation of the Z in both sexes

209 The previous patterns are consistent with an upregulation of the Z-chromosome in both sexes
210 after the degeneration of the W-specific region, and could represent the intermediate step in the
211 evolution of dosage compensation originally postulated by Ohno. However, they could also be
212 due to high expression of the ancestral proto-Z in both sexes, before sex chromosome
213 divergence. To exclude this, we identified one-to-one orthologs between genes annotated in
214 both species using a reciprocal best hit approach (7,382 orthologs, Figure 3-source data 1). All
215 genes that were classified as Z-specific in one species but as autosomal in the other were
216 considered to be part of the S1 strata (S1jap if they were Z-specific in *S. japonicum* or S1man if
217 they were Z-specific in *S. mansoni*). We then used the pseudoautosomal expression of these
218 lineage-specific strata as a proxy for the ancestral level of expression. For instance, in *S.*
219 *japonicum*, we estimated the S1jap:AA ratio, after normalizing the expression data by their
220 respective (pseudo)autosomal level in *S. mansoni* (Figure 3A and 3B). The reversed analysis
221 was performed for S1mans (Figure 3C and 3D).

222 Figure 3 confirms that the male-biased expression of Z-specific genes is a consequence of their
223 sex-linkage, and that the Z-chromosome has become under-expressed in females relative to the
224 ancestral expression. However, a full two-fold reduction in female expression is not observed,
225 consistent with partial upregulation, and/or full upregulation of a subset of dosage-sensitive
226 genes (Z:AA ranging from 0.68 to 0.83, Supplementary File 1)(48,49). Figure 3 also generally
227 supports an increase in expression in males (ZZ:AA ranging from 0.98 to 1.35, Supplementary
228 File 1). Male adults of *S. japonicum* are the exception, with a ZZ:AA of 0.98. However, given
229 that an excess of expression is observed when (i) we do not take into account the ancestral
230 expression (Figure 2), (ii) we focus on genes previously mapped to the ZW pair (Figure 2-figure
231 supplement 3, 4, 10 and 11), and (iiii) we consider the schistosomula stage (with or without the
232 ancestral expression and independent of the classification), this is likely due to noise in the
233 sample and not to a true biological difference (only 58 genes were tested). Figure 3 shows the

234 distributions for all genes with a minimum RPKM value of 1 in males and females of both
235 species. We repeated the analysis using the same filters as before (minimum RPKM of 0, TPM
236 of 0, TPM of 1), and with a publicly available list of 1:1 orthologs (obtained from the Wormbase
237 Biomart, see Methods). The resulting plots are shown in Figure 2-figure supplement 1 to Figure
238 2-figure supplement 14, and Figure 3-figure supplement 1 and 2. Gene expression values for
239 the orthologs of each species are provided in Figure 3-source data 2 and 3.
240
241 4. Male-biased protein dosage of Z-specific genes.
242 We tested for putative post-transcriptional mechanisms by assessing the dosage compensation
243 pattern at the proteomic level in adult *S. mansoni*, using a somatic tissue (head region) as well
244 as the gonads; three replicates were used for each tissue and sex. Heads and gonads were
245 chosen as they allowed us to compare Z-specific gene dosage in tissues with widespread
246 functional sex-specificity (ovary and testis), and in a tissue where most dosage imbalances are
247 likely to be deleterious. We used a label-free quantitative mass spectrometry approach to obtain
248 a relative quantification of the protein levels in each tissue depending of the sex (Figure 4-
249 source data 1 to 5). Post-transcriptional dosage compensation mechanisms would be
250 detectable by (i) an equalization of the Z expression between sexes at the protein level (F:M
251 close to 1 for both the Z and the autosomes); (ii) a different correlation between F:M obtained
252 from mRNA and from proteins for Z-linked and autosomal genes. We used publicly available
253 head and gonad microarray data (50) as the transcriptomic reference (Figure 4-source data 6).
254 A significant and positive correlation was found between the F:M ratio derived from the
255 microarray and from the proteomic data (Figure 4), and between transcript and protein dosage
256 levels in both males and females (Figure 4-figure supplement 1 and 2), confirming the validity of
257 the comparison.
258 Similar to what was observed using RNA-seq, the expression of Z-specific genes was strongly
259 male-biased compared to that of autosomal genes in both heads (F:M of 0.68 for the Z

chromosome versus 0.92 for the autosomes; Figure 4A, Supplementary File 1) and gonads (F:M of 0.78 versus 0.99; Figure 4B, Supplementary File 1). These F:M ratios are closer to each other than in our RNA-seq analysis (Figure 2, Supplementary File 1), or than the microarray data (Supplementary File 1), which could suggest a potential contribution of post-transcriptional regulation to dosage equalization. However, Figure 4B shows that Z-linked and autosomal genes show a similar correlation between the F:M ratios found for mRNAs and proteins ($p>0.05$ with a Fisher r-to-z transformation of the correlation coefficients, Figure 4C and 4D), which argues against a major role of post-transcriptional regulation to balance expression. This similarity between Z-linked and autosomal genes holds when only genes with male-biased expression in the microarray data are considered (Figure 4-figure supplement 3), and when the transcript and protein dosage of Z-linked autosomal genes are compared within each sex (Figure 4-figure supplement 1 and 2).

DISCUSSION

1. Schistosome sex chromosome evolution in the age of genomics

S. mansoni, *S. haematobium* and *S. japonicum* are the main species responsible for human schistosomiasis and have been the subject of many molecular and genomic studies. Despite the availability of extensive genomic and transcriptomic resources (e.g. a genome assembly at the near-chromosome level for *S. mansoni* (42,43), or sex- and stage-specific transcriptomes (46,51–53), many basic questions remain regarding their reproduction and biology. For instance, the master sex-determining gene (and whether it is located on the W or Z) is still a mystery (54,55). This is partly due to the inherent challenges of assembling genomes from sequencing data, especially for regions rich in heterochromatin and repetitive sequences, such as sex chromosomes. For instance, 416 scaffolds, including 3,893 genes (29% of the annotated nuclear genes), are still unplaced. By basing our analysis on genomic coverage, we were able to detect a further 285 Z-specific genes in *S. mansoni*; their role in sex determination can be

286 investigated further. Our comparative approach can also reduce the number of candidates, as
287 any gene involved in sex determination should in principle be found in the ancestral Z-specific
288 stratum; similar analyses in other species can in the future refine the candidate region. Another
289 advantage of basing our Z-assignment purely on coverage patterns is that our results should be
290 largely independent of potential biases in the current version of the genome. It should however
291 be noted that many genes are likely still missing from the current assembly (which has a
292 BUSCO score of 76% complete plus fragmented genes,
293 https://parasite.wormbase.org/index.html (56,57)), and that repeating these analyses using
294 future improved assemblies will be necessary to obtain the full set of sex-linked genes.
295 A gradient of ZW heteromorphism between schistosome species was revealed by cytogenetic
296 studies (39,58,59); in particular, African schistosomes were found to have much more extensive
297 ZW differentiation and W heterochromatinization than Asian species (60). Our results generally
298 support these cytogenetic data: we confirm the acquisition of independent evolutionary strata in
299 the sex chromosomes of *S. mansoni* and *S. japonicum*, and detect a larger number of Z-specific
300 genes in the African species (8% to 11% of all annotated orthologs, respectively, in *S. mansoni*
301 and *S. haematobium*) than in *S. japonicum* (5.5% of all annotated orthologs). Interestingly,
302 although the sex chromosomes of the African *S. mansoni* and *S. haematobium* differ
303 morphologically, they are largely similar in their gene content (Figure 1), consistent with their
304 much closer phylogenetic relationship (the median synonymous divergence between the two
305 species is around 17%, compared to 65% for *S. mansoni / S. japonicum*, Appendix 1). This may
306 be comparable to snakes, where ZW pairs with vastly different morphologies were all equally
307 differentiated at the genomic level (9), and highlights the contribution of other factors, such as
308 differential transposable element accumulation, to the large scale morphology of sex
309 chromosomes.
310
311 <u>2. ZW systems and incomplete dosage compensation: gene-by-gene or partial shift?</u>

ZW systems (aside from Lepidoptera) consistently show male-biased expression of the Z chromosome (Gu and Walters, 2017). While female-biased expression of the X occurs in a few young XY systems (16,31,57-61), well established X chromosomes generally show full equalization of gene expression between the sexes. This difference has often been framed as the acquisition of global mechanisms of dosage compensation, which affects the whole X/Z, versus the acquisition of local compensation, in which dosage-sensitive genes become individually regulated (61). Several parameters should influence this, and favor local compensation in ZW systems: (i) The speed of the heterochromosome degeneration: when only a few genes are lost at a time (because the region of suppressed recombination is small, or because degeneration is slow), the establishment of a gene-by-gene dosage compensation may be favored; on the other hand, the loss of many genes at once could favor global mechanisms of dosage compensation (23,62). Since more mutations occur during spermatogenesis than oogenesis, female-specific W chromosomes will generally have lower mutation and degeneration rates than male-specific Ys, favoring local compensation; (ii) The effective population size of Z (NeZ): NeZ is decreased when the variance in reproductive success of ZZ males is larger than that of ZW females (e.g. in the presence of strong sexual selection). This will impair the adaptive potential of the Z (63,64), such that only strongly dosage-sensitive genes can become upregulated in the heterogametic sex, while the others remain uncompensated; (iii) More efficient purging of mutations that are deleterious to males: strong sexual selection can also increase the strength of purifying selection on males, by preventing all but the fittest males from contributing to the next generations. If mutations that compensate for the loss of Y/W-linked genes overexpress the X/Z copy in both sexes, they will be under negative selection in the homogametic sex, and may be more efficiently selected against when males are the homogametic sex (64).

Schistosomes are unusual among female-heterogametic clades in that they appear to have a chromosome-wide upregulation of the Z in both sexes; such an increase in males was not

detected in birds (17) or snakes (9), even when ancestral expression was taken into account. They therefore likely represent an intermediate between ZW species with true local compensation, and the chromosome-wide compensation of the ZW Lepidoptera. These results further show that, even if mutations that upregulate gene expression in both sexes are more easily fixed on an evolving X-chromosome than on an evolving Z (64), this is not an absolute barrier to the evolution of global dosage compensation. It is however still unclear why the evolutionary dynamics appear to differ between schistosomes and most other ZW clades, as the demographic and population genetics parameters of this group are largely unknown. The observed male biased sex-ratio in adults, combined with a largely monogamous mating system (65,66), may increase the reproductive variance of males and could reduce the effective population size of the Z. This should also lead to stronger sexual selection in males than in females (66,67), suggesting similar evolutionary dynamics as in other ZW systems. A detailed characterization of the population genetics of the Z chromosome and autosomes will therefore be crucial for understanding what may have driven the evolution of this unusual system.

3. The relevance of the Ohno's hypothesis in the high-throughput sequencing era

Ohno's hypothesis predicts that the heterochromosome is initially overexpressed in both sexes, then secondarily downregulated in the homogametic sex (14),. This theoretical scenario was first formulated to account for the inactivation of the X in mammals. Since then, similar molecular mechanisms to downregulate the X/Z chromosome have been characterized in nematodes and moths (68,69). If an initial upregulation of the X did occur in both sexes, then inactivation in the homogametic sex should simply restore the ancestral expression levels, a hypothesis that has been tested in many empirical studies in mammals. Most of them assumed that the X and autosomes must have had similar ancestral levels of expression, and simply compared their expression (17,19,20,23,35,70–75). These yielded mixed results, with some (20) finding reduced expression of the X, while others (e.g. (71–73)) found similar levels of

expression for X-linked and autosomal genes, in agreement with Ohno's predictions. Taking ancestral gene expression into account, Julien et al. (17) found evidence of an Ohno-like mechanism in the marsupials but not in placental mammals (17). Pessia et al. (49) recently found that while individual dosage-sensitive genes do show evidence of up-regulation, the majority does not. The evolution of X-inactivation may therefore have involved a complex scenario under which a few dosage sensitive genes first became individually upregulated in both sexes (gene-by-gene compensation), followed by the establishment of a chromosome-wide mechanism to downregulate expression in females (global compensation) (48,49).

Our results, which consider ancestral expression and do not indicate a major influence of post-transcriptional regulation, suggest a scenario closer to Ohno's original hypothesis, with the male Z showing a consistent increase in expression. A similar pattern was observed in *Tribolium castaneum* (Coleoptera, (30)), where the female X was found to be over-expressed relative to the autosomes, and to the male X-chromosome. However, an RNA-seq analysis in the same species did not detect this (76), so it is at this point unclear whether it truly represents an example of Ohno's model in action. The youngest evolutionary stratum of the young XY pair of threespine sticklebacks also shows overexpression in females (31), even when ancestral expression is accounted for (77). However, the interpretation of these patterns is complicated by the fact that such an overexpression is also detected for the pseudoautosomal region, and that the oldest evolutionary stratum appears to lack dosage compensation altogether. Schistosomes may therefore not only represent an ideal system in which to investigate the evolution of dosage compensation in a ZW system, but also an unparalleled system for understanding the relevance of the model and predictions originally made by Ohno.

MATERIALS AND METHODS

A detailed description of the computational analyses, as well as all the scripts that were used, are provided in Appendix 1.

DNA sequencing of *S. japonicum* males and females

Male and female worms preserved in ethanol of *S. japonicum* were provided by Lu Dabing from Soochow University (Suzhou, China). DNA was extracted from 28 pooled males and 33 pooled females. The worms were lysed using the Tissue Lyser II kit (QIAGEN) and DNA was isolated using the DNeasy Blood & Tissue Kit (QIAGEN). DNA was then sheared with a AFA™ with Covaris Focused-ultrasonicator. Library preparation and sequencing (HiSeq 2500 v4 Illumina, 125bp paired-end reads) were performed at the Vienna Biocenter Next Generation sequencing facility (Austria). Reads have been deposited at the NCBI Short Reads Archive under accession number SRP135770.

Publicly available DNA reads and genome assemblies

S. mansoni DNA libraries (100bp paired-end reads) were downloaded from the NCBI Sequence Read Archive, under the accession numbers ERR562989 (~6,000 male pooled cercariae) and ERR562990 (~6,000 female pooled cercariae). Female *S. haematobium* DNA libraries (80bp paired-end reads) were found under accession numbers ERR037800 and ERR036251. No male *S. haematobium* library was available. The reference genome assemblies of *S. mansoni* (PRJEA36577, (78)), *S. haematobium* (PRJNA78265, (44)) and *S. japonicum* (PRJEA34885, (45)) were obtained from the WormBase parasite database (https://parasite.wormbase.org/index.html, (56,57)).

Orthology and assignment to the *S. mansoni* chromosomes

S. mansoni coding sequences and their respective chromosomal locations were obtained from the WormBase Parasite database (https://parasite.wormbase.org/index.html, (56,57)). This gene set was mapped to the *S. haematobium* and *S. japonicum* genome assemblies using Blat (79) with a translated query and dataset (-dnax option), and a minimum mapping score of 50; only the genome location with the best score was kept for each. When more than one gene overlapped by more than 20 base pairs, only the gene that had the highest mapping score was kept. Finally, each scaffold was assigned to one of the *S. mansoni* chromosomes, depending on the majority location of the genes that mapped to it, or on their total mapping scores if the same number of genes mapped to two separate chromosomes. The final chromosomal assignments are provided in Figure 1-source data 1.

DNA read mapping and estimation of genomic coverage

For the *S. japonicum* DNA reads, adaptors were removed using Cutadapt (v1.9.1, (80)) and the quality of the reads was assessed using FastQC (v0.11.2, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/); no further quality trimming was deemed necessary. For the *S. mansoni* and *S. haematobium* reads, potential adaptors were systematically removed, and reads were trimmed and filtered depending on their quality, using Trimmomatic (v0.36, (81)). The resulting read libraries of each species were mapped separately to their reference genomes using Bowtie2 (--end-to-end --sensitive mode, v2.2.9, (82)). The resulting alignments were filtered to keep only uniquely mapped reads, and the male and female coverages were estimated from the filtered SAM files with SOAPcoverage (v2.7.7., http://soap.genomics.org.cn/index.html). Coverage values were calculated for each scaffold in *S. haematobium* and *S. japonicum*, and for each 10kb non-overlapping window in *S. mansoni*. The coverage values for each library are provided in Figure 1-source data 3 and 4.

Detection of Z-specific sequences

441 For each species, we calculated the log2(female:male) coverage of each scaffold or, in the case
442 of *S. mansoni*, of each 10Kb window along the genome. Since only female DNA data was
443 available for *S. haematobium*, the log2(female1+female2) was used instead for this species.
444 In order to determine the 95% and 99% percentile of log2(female:male) of Z-linked sequences,
445 which we use as cut-off values for assignment to Z-specific regions, we first excluded
446 scaffolds/windows that fit an autosomal profile. To do so, the 1st and 5th percentile of
447 log2(female:male) were estimated using all 10kb windows found on the annotated autosomes of
448 *S. mansoni* (78); in *S. haematobium* and *S. japonicum*, all scaffolds that mapped to the *S.*
449 *mansoni* autosomes were used for this purpose.
450 By plotting the distribution of log2(Female:Male) for the scaffolds/windows that map to each
451 chromosome (Appendix 1), we determined that: (i) for *S. mansoni* the 1st percentile
452 (log2(female:male)=-0.26) discriminates effectively between autosomal and Z-specific windows;
453 (ii) for *S. haematobium* and *S. japonicum*, which have noisier coverage, the 1st percentile
454 included a significant fraction of Z-derived genes, and the 5st percentile was used instead
455 (respectively log2(Female)=4.57 and log2(female:male)=-0.40). All scaffolds with higher
456 log2(female:male) or log2(female) were excluded.
457 The 95th and 99th quantiles of coverage were then calculated for the remaining, putatively Z-
458 linked, sequences. By plotting all the coverage values along the *S. mansoni* Z-chromosome
459 (Appendix 1 and Figure 1), we determined that: (i) for *S. mansoni* and *S. haematobium*, the 95th
460 quantile (log2(female:male)=-0.4 and log2(female)=4.41, respectively) was an effective cut-off
461 for discriminating pseudoautosomal and Z-specific sequences; (ii) for *S. japonicum*, using the
462 95th percentile lead to the exclusion of many genes in Z-specific regions, and the 99st percentile
463 (log2(F:M)=-0.84) was used instead. The Z-specific or autosome assignation was finally
464 attributed as follows: (i) in *S. mansoni*, windows were classified as Z-linked if they displayed
465 log2(F:M)<-0.4 and as autosomal if log2(F:M)>-0.4; (ii) for *S. haematobium*, scaffolds with
466 log2(femalesum)<4.41 were classified as Z-linked, scaffolds with log2(femalesum)>4.57 as

autosomal, and all others as ambiguous; (iii) for *S. japonicum*, scaffolds with log2(F:M)<-0.84 were classified as Z-linked, scaffolds with log2(F:M)>-0.40 as autosomal, and all others as ambiguous. For *S. haematobium* and *S. japonicum*, we considered only scaffolds with at least a coverage of 1 in each library (n.a.). In *S. mansoni*, unplaced scaffolds shorter than 50kb were excluded; five consecutive 10kb windows with consistent coverage patterns were required for a region to be classified as either Z-specific or autosomal. Smaller regions, as well as the two 10Kb windows surrounding them, were excluded (see Appendix 1). The final classifications for the three species are provided in Figure 1-source data 2, 3, and 4, and summarized for orthologs in Figure 1-source data 1.

Definition of Z-specific strata in *S. mansoni* and *S. japonicum*

The Z-specific gene content of *S. mansoni* and *S. japonicum* was compared in order to define Z-chromosome strata. Genes that were located on Z-specific scaffolds/windows in both species were assigned to the shared stratum "S0". Genes that were assigned to Z-specific regions in one species but not in the other were assigned to lineage-specific strata: “S1mans” genes were Z-specific in *S. mansoni* and autosomal in *S. japonicum*, while “S1jap” genes were Z-specific in *S. japonicum* and autosomal in *S. mansoni*. While the main figures consider all the genes that were classified as Z-linked or autosomal based on coverage (referred to as the "exhaustive classification" in Appendix 1 and Figures), independent of their original genomic location, we repeated the analyses using only the Z-specific genes that were already assigned to the ZW linkage map of *S. mansoni* (the "stringent classification" in Appendix 1). All genes belonging to the categories “excluded”, “ambiguous”, “n.a.” or that did not have orthologs on *S. japonicum* scaffolds were not further considered (Table 1). “PSA_shared” and “Aut_shared” are common to the two classifications and correspond to genes that were classified as autosomal in both species using coverage, and that were previously mapped to the ZW linkage group or to the autosomes of *S. mansoni*, respectively (78).

493

494 **Publicly available RNA reads and estimation of gene expression**

495 *S. mansoni* and *S. japonicum* RNA-seq libraries were obtained from SRA (NCBI). Accession
496 numbers are: *S. mansoni* adult females: ERR506076, ERR506083, ERR506084; *S. mansoni*
497 adult males: ERR506088, ERR506082, ERR506090; *S. mansoni* schistosomula females:
498 SRR3223443, SRR3223444; *S. mansoni* schistosomula males: SRR3223428, SRR3223429 ;
499 *S.japonicum* adult females: SRR4296944, SRR4296942, SRR4296940; *S.japonicum* adult
500 males: SRR4296945, SRR4296943, SRR4296941; *S. japonicum* schistosomula females:
501 SRR4279833, SRR4279491, SRR4267990; *S. japonicum* schistosomula males: SRR4279840,
502 SRR4279496, SRR4267991. Raw reads were cleaned using trimmomatic (v 0.36,(81)), and the
503 quality of the resulting reads was assessed using FastQC (v0.11.2,
504 https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Reads were mapped to their
505 respective reference genomes used Tophat2 (83). Read counts were obtained with Htseq (84)
506 and expression values (in Reads Per Kilobase of transcript per Millionmapped reads, RPKM)
507 were calculated for each gene in each of the RNA-seq libraries (Figure 2-source data 1). TPM
508 (Transcripts Per Kilobase Million) values were also calculated using Kallisto (85) against set of
509 coding sequences of the respective species. All expression values are provided in Figure 3-
510 source data 2 and 3. A Loess Normalization (R library Affy) was performed on the
511 schistosomulum data and on the adult data separately and all analyses were performed using
512 different thresholds (RPKM>0, RPKM>1, TPM>0 or TPM>1 in all the libraries of the studied
513 stage). The Loess normalization was applied to all conditions at once when we filtered for
514 minimum expression in all stages and sexes (RPKM>1 and RPKM>3, Figure 2-figure
515 supplement 15 -17, Figure 3-figure supplement 1 and 2). Correlation analyses were performed
516 for each developmental stage, considering libraries from both males and females, and both
517 species. As shown in Appendix 1, two *S. mansoni* libraries (ERR506076 and ERR506082) were
518 not well correlated with the other samples, and were excluded from our study. Expression

519 values were averaged for each stage and sex. The significance of differences between medians
520 of expression was tested with Wilcoxon rank sum tests with continuity correction.
521
522 **Detection of *S. mansoni* and *S. japonicum* one-to-one orthologs**
523 *S. japonicum* coding DNA sequences and their respective location on the genome scaffolds
524 were obtained from the WormBase Parasite database
525 (https://parasite.wormbase.org/index.html, (56,57)). The *S. mansoni* set of coding sequences
526 (see above) was mapped to the *S. japonicum* gene set using Blat (79) with a translated query
527 and dataset (-dnax option), and a minimum mapping score of 50; only reciprocal best hits were
528 kept. This reciprocal best hit ortholog list is provided in Figure 3-source data 1 and 2. A second
529 list of orthologs was obtained from the Biomart of WormBase Parasite, excluding paralogues,
530 and requiring a gene stable ID for both *S. mansoni* (PRJEA36577) and *S. japonicum*
531 (PRJEA34885) (in Figure 3-source data 1 and 3). In subsequent transcriptomic analyses each
532 list was used independently to ensure that the results were independent of the method used to
533 assign orthology.
534
535 **Microarray analysis**
536 Microarray data for male and female heads and gonads (50) were obtained from the Gene
537 Expression Omnibus (GEO) database (NCBI,
538 ftp://ftp.ncbi.nlm.nih.gov/geo/series/GSE23nnn/GSE23942/matrix/). A Loess Normalization (R
539 library Affy) was performed on the head and gonad data separately. When different probes
540 corresponded to one gene, their expression values were averaged. Gene expression was
541 available for a total of 6,925 genes. The normalized data are available in Figure 4-source data
542 6.
543 **Protein extraction and proteomic analysis**

544 Male and female adult *S. mansoni* gonads were sampled using the whole-organ isolation
545 approach described previously (86). Twenty ovaries and 20 testes, as well as five heads of each
546 sex, were sampled, in triplicate, from paired worms. All biological samples were resuspended in
547 Laemmli buffer, denatured and frozen at -20°C until further processing. Subsequent protein
548 treatment and analyses were performed at the "EDyP-service" – proteomic platform (Grenoble,
549 France). The extracted proteins were digested by modified trypsin (Promega, sequencing
550 grade). The resulting peptides were analyzed by nanoLC-MS/MS (Ultimate 3000 RSLCnano
551 system coupled to Q-Exactive Plus, both Thermo Fisher Scientific). Separation was performed
552 on a 75µm x 250mm C18 column (ReproSil-Pur 120 C18-AQ 1.9µm, Dr. Maisch GmbH) after a
553 pre-concentration and desalting step on a 300 µm × 5 mm C18 precolumn (Pepmap, Thermo
554 Fisher Scientific).
555 MS and MS^2 data were acquired using Xcalibur (Thermo Fisher Scientific). Full-scan (MS)
556 spectra were obtained from 400 to 1600m/z at a 70.000 resolution (200m/z). For each full-scan,
557 the most intense ions (top 10) were fragmented in MS^2 using high-energy collisional dissociation
558 (HCD). The obtained data were processed in MaxQuant 1.5.8.3 against the database loaded
559 from Uniprot (taxonomy *Schistosoma mansoni*, October 26th, 2017, 13.521 entries) and the
560 MaxQuant embedded database of frequently observed contaminants. The resulted iBAQ values
561 (87) were loaded into ProStaR (88) for statistical analysis. Contaminant and reverse proteins
562 were removed and only the proteins with three quantified values in at least one condition were
563 taken into account.
564 After log2 transformation, the iBAQ values were normalized by overall-wise median centering
565 followed by imputation using detQuantile algorithm with quantile set to 1 (Figure 4-source data 1
566 and 2). An alternative set of data without imputed values is available in Figure 4-source data 3
567 and 4. 1,988 and 2,750 *Schistosoma mansoni* proteins were identified in heads and in gonads,
568 respectively (See Figure 4-source data 1 and 2 for statistical testing of differential abundance
569 between male and female samples). Among them, 1,741 and 2,516 could be attributed

570 unambiguously to a *Schistosoma mansoni* gene, and were represented by more than one
571 peptide; these were subsequently analyzed (See Figure 4-source data 5).
572
573 **Data availability**
574 DNA reads of male and female *S. japonicum* will be released upon publication on the SRA
575 database under study number SRP135770. Sex and tissue specific *S. mansoni* label-free
576 proteomic data are provided in Figure 4-source data 1 to 4.
577
578 **Code availability**
579 The full bioinformatic pipeline used in this study is provided in Appendix 1.
580
581 **ACKNOWLEDGMENTS**
582 We are grateful to Lu Dabing (Soochow University, Suzhou, China) for providing *Schistosoma*
583 *japonicum* samples, to Ariana Macon (IST Austria) and Georgette
584 Stovall (JLU Giessen) for technical assistance, to IT support at IST Austria for providing optimal
585 environment to bioinformatic analyses, and to the Vicoso lab for comments on the manuscript.
586 We thank the support of the discovery platform and informatics group at EDyP. Proteomic
587 experiments were partly supported by the Proteomics French Infrastructure (ANR-10-INBS-08-
588 01 grant) and Labex GRAL (ANR-10-LABX-49-01); and the Wellcome Trust, grant
589 107475/Z/15/Z to CG.G. and T.Q. (FUGI). This project was funded by an Austrian Science
590 Foundation FWF grant (Project P28842) to B.V.
591
592 **CONFLICT OF INTEREST STATEMENT**
593 The authors declare that they have no conflict of interest.
594
595 **LIST OF FIGURES**

LIST OF TABLES

SUPPLEMENTARY FILES

Supplementary File 1. Comparison of the ratio of expression between the Z-linked and autosomal genes (Z:A), and comparison of expression between males and females (F:M) in the different species, stages and methods (Supplementary Table 1). Comparison of female:male ratios of expression (F:M) using microarray and proteomics (Supplementary Table 2)

Appendix 1: Full description of the bioinformatics pipelines used for each analysis**.**

SUPPLEMENTARY FIGURES

Figure 2 - figure supplements 2 to 14. Adult/Schistosomulum expression patterns (in RPKM/TPM) of genes located in the different strata of the Z (S0, S1man and S1jap), as well as pseudoautosomal and autosomal genes.

Figure 2 - figure supplement 15. Z-linked and autosomal female:male ratio of gene expression using different filters.

Figure 2 - figure supplement 16. Z-linked and autosomal gene expression in females and males using different filters.

Figure 2 - figure supplement 17. Z-linked and autosomal female:male ratio of gene expression according to presence/absence of known protein-protein interactions.

Figure 3 - figure supplement 1. Z-linked and autosomal female:male ratio of gene expression, normalized by ancestral autosomal expression, and using different filters.

Figure 3 - figure supplement 2. Z-linked and autosomal gene expression in females and males, normalized by ancestral autosomal expression, and using different filters.

Figure 4 - figure supplement 1. Pearson correlations between gene dosage at the transcript and protein levels in male heads and gonads.

Figure 4 - figure supplement 2. Pearson correlations between gene dosage at the transcript and protein levels in female heads and gonads.

Figure 4 - figure supplement 3. Pearson correlations between the female:male of expression obtained by proteomics (y-axes) and by microarrays (x-axes) in *S.mansoni* heads and gonads, using only genes with male-biased expression in the microarray data.

SUPPLEMENTARY DATASETS

Figure 1 - source data 1. Comparative genomics: coverage analysis and strata identification

Figure 1 - source data 2. *S. mansoni* reference species: coverage analysis and Z-Autosome assignment

Figure 1 - source data 3. *S. haematobium* autosome vs Z-specific region assigment

Figure 1 - source data 4. *S. japonicum* autosome vs Z-specific region assigment

Figure 2 - source data 1. RPKM calculation for *S. mansoni* gene expression

Figure 2 - source data 2. RPKM calculation for *S. japonicum* gene expression

Figure 3 - source data 1. One-to-one orthology *S. mansoni* vs *S. japonicum*

Figure 3 - source data 2. Transcriptomic data, for blat 1-to-1 orthologs

Figure 3 - source data 3. Transcriptomic data, for WormBase Biomart orthologs

Figure 4 - source data 1. Proteomic data with imputed values – GONADS

Figure 4 - source data 2. Proteomic data with imputed values – HEADS

Figure 4 - source data 3. Proteomic data without imputed values – GONADS

Figure 4 - source data 4. Proteomic data without imputed values – HEADS

Figure 4 - source data 5. Correspondance between Gene_Id and Protein_ID

Figure 4 - source data 6. Microarray data

REFERENCES

1. Bachtrog D, Mank JE, Peichel CL, Kirkpatrick M, Otto SP, Ashman T-L, et al. Sex Determination: Why So Many Ways of Doing It? PLoS Biol. 2014;12(7):e1001899.
2. Charlesworth B. The evolution of sex chromosomes. Science. 1991;251(4997).
3. Charlesworth D, Charlesworth B, Marais G. Steps in the evolution of heteromorphic sex chromosomes. Heredity. 2005;95(2):118–28.
4. Rice WR. The Accumulation of Sexually Antagonistic Genes as a Selective Agent Promoting the Evolution of Reduced Recombination between Primitive Sex Chromosomes. Evolution. 1987;41(4):911.
5. Bergero R, Charlesworth D. The evolution of restricted recombination in sex chromosomes. Trends Ecol Evol. 2009;24(2):94–102.
6. Engelstadter J. Muller's Ratchet and the Degeneration of Y Chromosomes: A Simulation Study. Genetics. 2008;180(2):957–67.
7. Ellegren H. Sex-chromosome evolution: recent progress and the influence of male and female heterogamety. Nat Rev Genet. 2011;12(3):157–66.
8. Wang J, Na J-K, Yu Q, Gschwend AR, Han J, Zeng F, et al. Sequencing papaya X and Yh chromosomes reveals molecular basis of incipient sex chromosome evolution. Proc Natl Acad Sci. 2012;109(34):13710–5.
9. Vicoso B, Emerson JJ, Zektser Y, Mahajan S, Bachtrog D. Comparative Sex Chromosome Genomics in Snakes: Differentiation, Evolutionary Strata, and Lack of Global Dosage Compensation. PLoS Biol. 2013:e1001643.
10. Vicoso B, Kaiser VB, Bachtrog D. Sex-biased gene expression at homomorphic sex chromosomes in emus and its implication for sex chromosome evolution. Proc Natl Acad Sci. 2013;110(16):6453–8.
11. Lemaitre C, Braga MD V, Gautier C, Sagot M-F, Tannier E, Marais GAB. Footprints of inversions at present and past pseudoautosomal boundaries in human sex chromosomes. Genome Biol Evol. 2009;1(0):56–66.
12. Guo M, Davis D, Birchler JA. Dosage effects on gene expression in a maize ploidy series. Genetics. 1996;142(4):1349–55.
13. Wijchers PJ, Festenstein RJ. Epigenetic regulation of autosomal gene expression by sex chromosomes. Trends Genet. 2011;27(4):132–40.
14. Ohno S. Sex Chromosomes and Sex-linked Genes. Springer-Verlag. 1967.
15. Gartler SM. A brief history of dosage compensation. J Genet. 2014;93(2):591–5.
16. Gu L, Walters JR. Evolution of Sex Chromosome Dosage Compensation in Animals: A Beautiful Theory, Undermined by Facts and Bedeviled by Details. Genome Biol Evol. 2017;9(9):2461–76.
17. Julien P, Brawand D, Soumillon M, Necsulea A, Liechti A, Schütz F, et al. Mechanisms and Evolutionary Patterns of Mammalian and Avian Dosage Compensation. PLoS Biol. 2012;10(5):e1001328.
18. Lin F, Xing K, Zhang J, He X. Expression reduction in mammalian X chromosome evolution refutes Ohno's hypothesis of dosage compensation. Proc Natl Acad Sci. 2012;109(29):11752–7.
19. Nguyen DK, Disteche CM. Dosage compensation of the active X chromosome in

mammals. Nat Genet. 2006;38(1):47–53.
20. Xiong Y, Chen X, Chen Z, Wang X, Shi S, Wang X, et al. RNA sequencing shows no dosage compensation of the active X-chromosome. Nat Genet. 2010;42(12):1043–7.
21. Vicoso B, Bachtrog D. Progress and prospects toward our understanding of the evolution of dosage compensation. Chromosome Res. 2009;17(5):585–602.
22. Mank JE. Sex chromosome dosage compensation: definitely not for everyone. Trends Genet. 2013;29(12):677–83.
23. Gu L, Walters JR. Evolution of Sex Chromosome Dosage Compensation in Animals: A Beautiful Theory, Undermined by Facts and Bedeviled by Details. Genome Biol Evol. 2017;9(9):2461–76.
24. Itoh Y, Melamed E, Yang X, Kampf K, Wang S, Yehya N, et al. Dosage compensation is less effective in birds than in mammals. J Biol. 2007;6(1):2.
25. Ellegren H, Hultin-Rosenberg L, Brunström B, Dencker L, Kultima K, Scholz B. Faced with inequality: chicken do not have a general dosage compensation of sex-linked genes. BMC Biol. 2007;5(1):40.
26. Arnold AP, Itoh Y, Melamed E. A Bird's-Eye View of Sex Chromosome Dosage Compensation. Annu Rev Genomics Hum Genet. 2008;9(1):109–27.
27. Wolf JB, Bryk J. General lack of global dosage compensation in ZZ/ZW systems? Broadening the perspective with RNA-seq. BMC Genomics. 2011;12(1):91.
28. Chen S, Zhang G, Shao C, Huang Q, Liu G, Zhang P, et al. Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle. Nat Genet. 2014;46(3):253–60.
29. Huylmans AK, Macon A, Vicoso B. Global Dosage Compensation Is Ubiquitous in Lepidoptera, but Counteracted by the Masculinization of the Z Chromosome. Mol Biol Evol. 2017;34(10):2637–49.
30. Prince EG, Kirkland D, Demuth JP. Hyperexpression of the X Chromosome in Both Sexes Results in Extensive Female Bias of X-Linked Genes in the Flour Beetle. Genome Biol Evol. 2010;2(0):336–46.
31. Schultheiß R, Viitaniemi HM, Leder EH. Spatial Dynamics of Evolving Dosage Compensation in a Young Sex Chromosome System. Genome Biol Evol. 2015;7(2):581–90.
32. Vicoso B, Bachtrog D. Numerous Transitions of Sex Chromosomes in Diptera. PLOS Biol. 2015;13(4):e1002078.
33. Gu L, Walters JR, Knipple DC. Conserved Patterns of Sex Chromosome Dosage Compensation in the Lepidoptera (WZ/ZZ): Insights from a Moth Neo-Z Chromosome. Genome Biol Evol. 2017;9(3):802–16.
34. Uebbing S, Konzer A, Xu L, Backström N, Brunström B, Bergquist J, et al. Quantitative Mass Spectrometry Reveals Partial Translational Regulation for Dosage Compensation in Chicken. Mol Biol Evol. 2015;32(10):2716–25.
35. Chen X, Zhang J. No X-Chromosome Dosage Compensation in Human Proteomes. Mol Biol Evol. 2015;32(6):1456–60.
36. Chitsulo L, Loverde P, Engels D. Schistosomiasis. Nat Rev Microbiol. 2004;2(1):12–3.
37. Loker ES, Brant S V. Diversification, dioecy and dimorphism in schistosomes. Trends Parasitol. 2006;22(11):521–8.

744 38. Kunz W. Schistosome male-female interaction: induction of germ-cell differentiation.
745 Trends Parasitol. 2001;17(5):227–31.
746 39. Grossman AI, Short RB, Cain GD. Karyotype evolution and sex chromosome
747 differentiation in Schistosomes (Trematoda, Schistosomatidae). Chromosoma.
748 1981;84(3):413–30.
749 40. Lawton SP, Hirai H, Ironside JE, Johnston DA, Rollinson D. Genomes and geography:
750 genomic insights into the evolution and phylogeography of the genus Schistosoma.
751 Parasit Vectors. 2011;4(1):131.
752 41. Vicoso B, Bachtrog D. Lack of Global Dosage Compensation in Schistosoma mansoni, a
753 Female-Heterogametic Parasite. Genome Biol Evol. 2011;3(0):230–5.
754 42. Protasio A V, Tsai IJ, Babbage A, Nichol S, Hunt M, Aslett MA, et al. A systematically
755 improved high quality genome and transcriptome of the human blood fluke Schistosoma
756 mansoni. PLoS Negl Trop Dis. 2012;6(1):e1455.
757 43. Criscione CD, Valentim CL, Hirai H, LoVerde PT, Anderson TJ. Genomic linkage map of
758 the human blood fluke Schistosoma mansoni. Genome Biol. 2009;10(6):R71.
759 44. Young ND, Jex AR, Li B, Liu S, Yang L, Xiong Z, et al. Whole-genome sequence of
760 Schistosoma haematobium. Nat Genet. 2012;44(2):221–5.
761 45. Zhou Y, Zheng H, Chen Y, Zhang L, Wang K, Guo J, et al. The Schistosoma japonicum
762 genome reveals features of host–parasite interplay. Nature. 2009;460(7253):345–51.
763 46. Picard MAL, Boissier J, Roquis D, Grunau C, Allienne J-F, Duval D, et al. Sex-Biased
764 Transcriptome of Schistosoma mansoni: Host-Parasite Interaction, Genetic Determinants
765 and Epigenetic Regulators Are Associated with Sexual Differentiation. PLoS Negl Trop
766 Dis. 2016;10(9):e0004930.
767 47. Wang J, Yu Y, Shen H, Qing T, Zheng Y, Li Q, et al. Dynamic transcriptomes identify
768 biogenic amines and insect-like hormonal regulation for mediating reproduction in
769 Schistosoma japonicum. Nat Commun. 2017;8:14693.
770 48. Sangrithi MN, Royo H, Mahadevaiah SK, Ojarikre O, Bhaw L, Sesay A, et al. Non-
771 Canonical and Sexually Dimorphic X Dosage Compensation States in the Mouse and
772 Human Germline. Dev Cell. 2017;40(3):289–301.e3.
773 49. Pessia E, Makino T, Bailly-Bechet M, McLysaght A, Marais GAB. Mammalian X
774 chromosome inactivation evolved as a dosage-compensation mechanism for dosage-
775 sensitive genes on the X chromosome. Proc Natl Acad Sci. 2012;109(14):5346–51.
776 50. Nawaratna SSK, McManus DP, Moertel L, Gobert GN, Jones MK. Gene Atlasing of
777 digestive and reproductive tissues in Schistosoma mansoni. PLoS Negl Trop Dis.
778 2011;5(4):e1043.
779 51. Lu Z, Sessler F, Holroyd N, Hahnel S, Quack T, Berriman M, et al. Schistosome sex
780 matters: a deep view into gonad-specific and pairing-dependent transcriptomes reveals a
781 complex gender interplay. Sci Rep. 2016;6:31150.
782 52. Grevelding CG, Langner S, Dissous C. Kinases: Molecular Stage Directors for
783 Schistosome Development and Differentiation. Trends Parasitol. 2017.
784 53. Lu Z, Sessler F, Holroyd N, Hahnel S, Quack T, Berriman M, et al. A gene expression
785 atlas of adult Schistosoma mansoni and their gonads. Sci Data. 2017;4:170118.
786 54. Lepesant JM, Cosseau C, Boissier J, Freitag M, Portela J, Climent D, et al. Chromatin
787 structural changes around satellite repeats on the female sex chromosome in

Schistosoma mansoni and their possible role in sex chromosome emergence. Genome Biol. 2012;13(2):R14.
55. Portela J, Grunau C, Cosseau C, Beltran S, Dantec C, Parrinello H, et al. Whole-genome in-silico subtractive hybridization (WISH) - using massive sequencing for the identification of unique and repetitive sex-specific sequences: the example of Schistosoma mansoni. BMC Genomics. 2010;11(1):387.
56. Howe KL, Bolt BJ, Cain S, Chan J, Chen WJ, Davis P, et al. WormBase 2016: expanding to enable helminth genomic research. Nucleic Acids Res. 2016;44(D1):D774–80.
57. Howe KL, Bolt BJ, Shafie M, Kersey P, Berriman M. WormBase ParaSite − a comprehensive resource for helminth genomics. Mol Biochem Parasitol. 2017;215:2–10.
58. Grossman AI, Short RB, Kuntz RE. Somatic chromosomes of Schistosoma rodhaini, S. mattheei, and S. intercalatum. J Parasitol. 1981;67(1):41–4.
59. Short RB, Grossman AI. Conventional giemsa and C-banded karyotypes of Schistosoma mansoni and S. rodhaini. J Parasitol. 1981;67(5):661–71.
60. Hirai H, Taguchi T, Saitoh Y, Kawanaka M, Sugiyama H, Habe S, et al. Chromosomal differentiation of the Schistosoma japonicum complex. Int J Parasitol. 2000;30(4):441–52.
61. Mank JE, Ellegren H. All dosage compensation is local: Gene-by-gene regulation of sex-biased expression on the chicken Z chromosome. Heredity. 2009;102(3):312–20.
62. Vicoso B, Charlesworth B. Effective population size and the faster-x effect: an extended model. Evolution. 2009;63(9):2413–26.
63. Mank JE. The W, X, Y and Z of sex-chromosome dosage compensation. Trends Genet. 2009;25(5):226–33.
64. Mullon C, Wright AE, Reuter M, Pomiankowski A, Mank JE. Evolution of dosage compensation under sexual selection differs between X and Z chromosomes. Nat Commun. 2015;6(1):7720.
65. Beltran S, Boissier J. Are schistosomes socially and genetically monogamous? Parasitol Res. 2009;104(2):481–3.
66. Beltran S, Boissier J. Schistosome monogamy: who, how, and why? Trends Parasitol. 2008;24(9):386–91.
67. Steinauer ML, Hanelt B, Agola LE, Mkoji GM, Loker ES. Genetic structure of Schistosoma mansoni in western Kenya: The effects of geography and host sharing. Int J Parasitol. 2009;39(12):1353–62.
68. Kiuchi T, Koga H, Kawamoto M, Shoji K, Sakai H, Arai Y, et al. A single female-specific piRNA is the primary determiner of sex in the silkworm. Nature. 2014;509(7502):633–6.
69. Meyer BJ. Targeting X chromosomes for repression. Curr Opin Genet Dev. 2010;20(2):179–89.
70. Graves JAM. Evolution of vertebrate sex chromosomes and dosage compensation. Nat Rev Genet. 2016;17(1):33–46.
71. Deng X, Hiatt JB, Nguyen DK, Ercan S, Sturgill D, Hillier LW, et al. Evidence for compensatory upregulation of expressed X-linked genes in mammals, Caenorhabditis elegans and Drosophila melanogaster. Nat Genet. 2011;43(12):1179–85.
72. Kharchenko P V, Xi R, Park PJ. Evidence for dosage compensation between the X chromosome and autosomes in mammals. Nat Genet. 2011;43(12):1167–9.
73. Yildirim E, Sadreyev RI, Pinter SF, Lee JT. X-chromosome hyperactivation in mammals

via nonlinear relationships between chromatin states and transcription. Nat Struct Mol Biol. 2011;19(1):56–61.
74. Lin F, Xing K, Zhang J, He X. Expression reduction in mammalian X chromosome evolution refutes Ohno's hypothesis of dosage compensation. Proc Natl Acad Sci. 2012;109(29):11752–7.
75. Pessia E, Engelstädter J, Marais GAB. The evolution of X chromosome inactivation in mammals: the demise of Ohno's hypothesis? Cell Mol Life Sci. 2014;71(8):1383–94.
76. Mahajan S, Bachtrog D. Sex-specific adaptation drives early sex chromosome evolution in Drosophila. Genome Biol Evol. 2015;337(2):341–5.
77. White MA, Kitano J, Peichel CL. Purifying Selection Maintains Dosage-Sensitive Genes during Degeneration of the Threespine Stickleback Y Chromosome. Mol Biol Evol. 2015;32(8):1981–95.
78. Protasio A V., Tsai IJ, Babbage A, Nichol S, Hunt M, Aslett MA, et al. A Systematically Improved High Quality Genome and Transcriptome of the Human Blood Fluke Schistosoma mansoni. PLoS Negl Trop Dis. 2012;6(1):e1455.
79. Kent WJ. BLAT--the BLAST-like alignment tool. Genome Res. 2002;12(4):656–64.
80. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011;17(1):10.
81. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20.
82. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357–9.
83. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 2009;25(9):1105–11.
84. Anders S, Pyl PT, Huber W. HTSeq–A Python framework to work with high-throughput sequencing data. Bioinformatics. 2014;btu638.
85. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol. 2016;34(5):525–7.
86. Hahnel S, Lu Z, Wilson RA, Grevelding CG, Quack T. Whole-Organ Isolation Approach as a Basis for Tissue-Specific Analyses in Schistosoma mansoni. PLoS Negl Trop Dis. 2013;7(7):e2336.
87. Tyanova S, Temu T, Cox J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. Nat Protoc. 2016;11(12):2301–19.
88. Wieczorek S, Combes F, Lazar C, Giai Gianetto Q, Gatto L, Dorffer A, et al. DAPAR & ProStaR: software to perform statistical analyses in quantitative discovery proteomics. Bioinformatics. 2017;33(1):135–6.

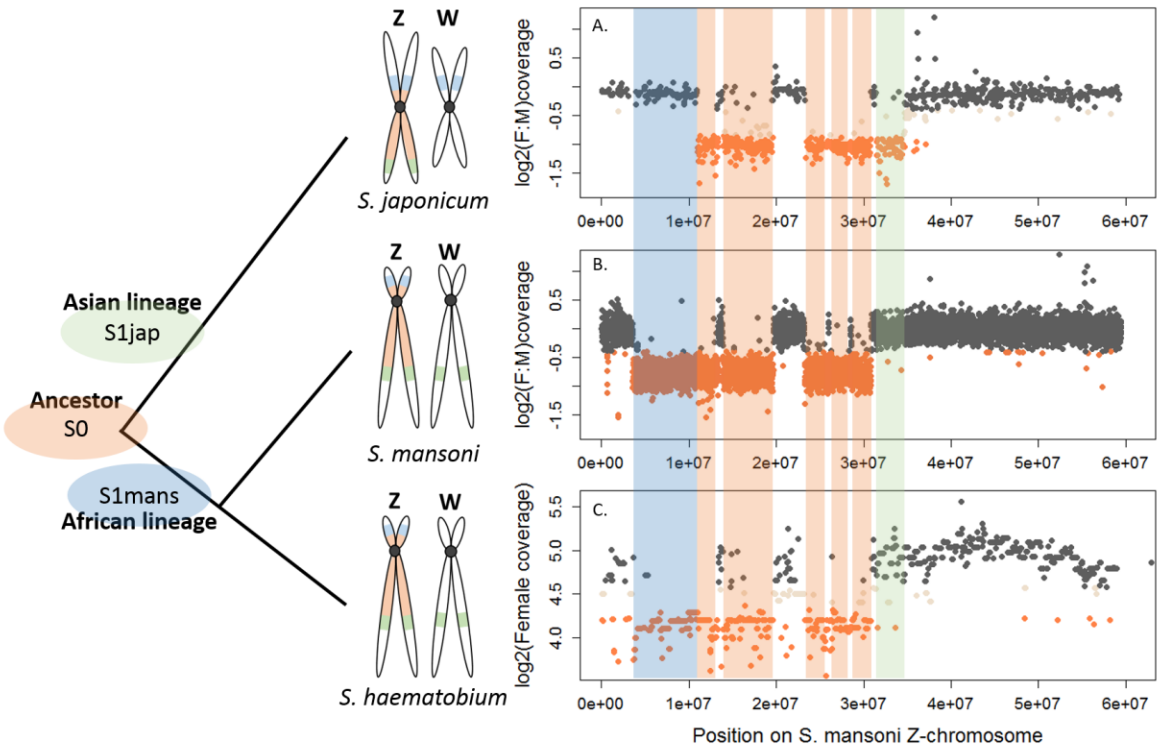


Figure 1. Shared and lineage-specific evolutionary strata on the Z-chromosome. The phylogeny of the three species is represented on the left. The Female:Male (F:M) ratio of coverage (y-axis) along the Z-chromosome of *S. mansoni* (x-axis) is shown for *S. japonicum* scaffolds **(A)** and *S. mansoni* 10kb windows **(B)**. Female coverage is shown for *S. haematobium* scaffolds **(C)**. All species share an ancestral Z-linked stratum S0 (marked in orange). The stratum S1jap (in green) is specific to the Asian lineage represented by *S. japonicum*. The stratum S1mans (in blue) is specific to the African lineage, represented by *S. mansoni* and *S. haematobium*. Dot color is attributed depending on the window/scaffold location within each species: Z-specific regions in orange, pseudoautosomal regions in grey, and ambiguous regions in beige.

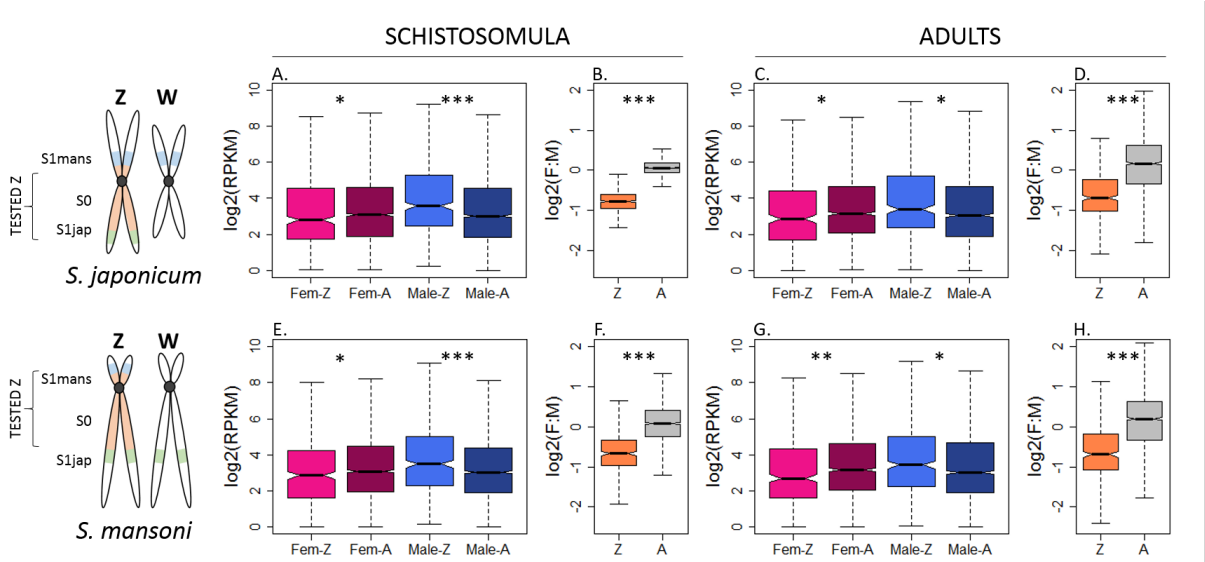


Figure 2. Patterns of expression on the Z and autosomes of *S. japonicum* and *S. mansoni* species.

Z-linked and autosomal gene expression patterns are shown for *S. japonicum* **(A-D)** and *S. mansoni* **(E-H)**, in non-differentiated schistosomula and sexually mature adults. In panels A, C, E, and G, Fem-Z and Male-Z refer to the expression of Z-linked genes in females and males, respectively, and Fem-A and Male-A to the expression of the autosomal genes in females and males. In panels B, D, F, and H, Z refers to Z-linked genes and A to autosomal genes. In all conditions, a strong male bias is observed for the Z-linked genes (**B**, **D**, **F**, **H**). This male-biased expression of the Z-linked genes is accompanied by both an under-expression in females and an over-expression in males, compared to the level of autosomal expression (**A**, **C**, **E**, **G**). The level of significance of each comparison (Wilcoxon rank sum test with continuity correction) is indicated by asterisks: **P-value*<0.05, ***P-value*<0.001, ****P-value*<0.0001.

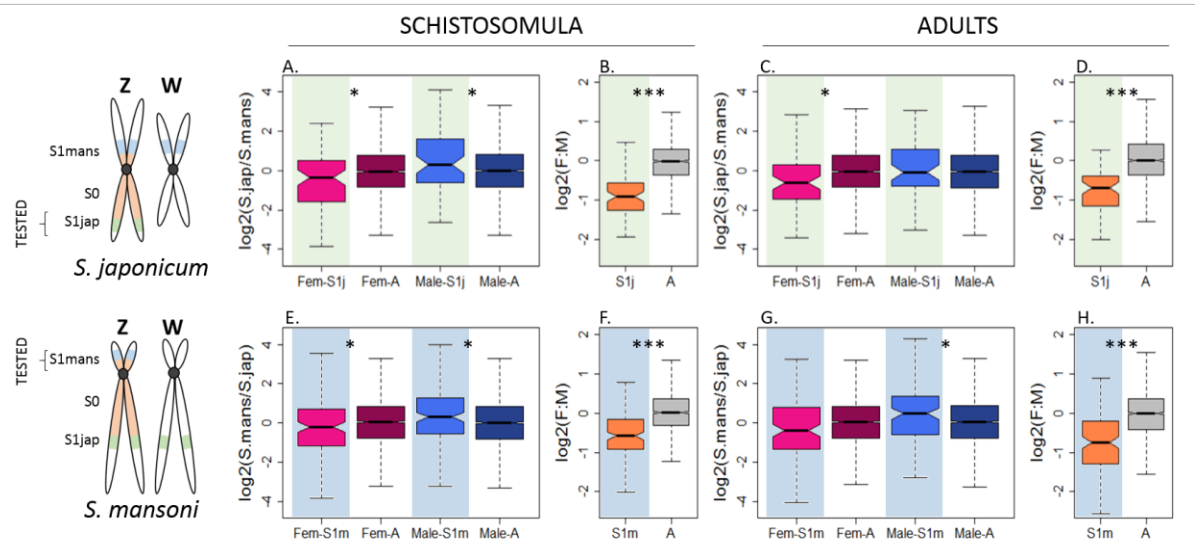


Figure 3. Convergent changes in the expression of Z-linked genes in *S. japonicum* and *S. mansoni* after sex chromosome differentiation.

The Z-linked and autosomal gene expression patterns (normalized by ancestral pseudoautosomal or autosomal expression, to show changes since the appearance of the two S1 strata) are shown for S1jap in *S. japonicum* **(A-D)** and S1mans in *S. mansoni* **(E-F)**, Fem-S1j and Male-S1j refer to the normalized expression levels of genes in the stratum S1jap in females and males, Fem-S1m and Male-S1m refer to the normalized expression levels in S1mans, and Fem-A and Male-A refer to the normalized expression levels of autosomal genes in females and in males, respectively. The level of significance of each comparison (Wilcoxon rank sum test with continuity correction) is denoted with asterisks: **P-value*<0.05, ***P-value*<0.001, ****P-value*<0.0001.

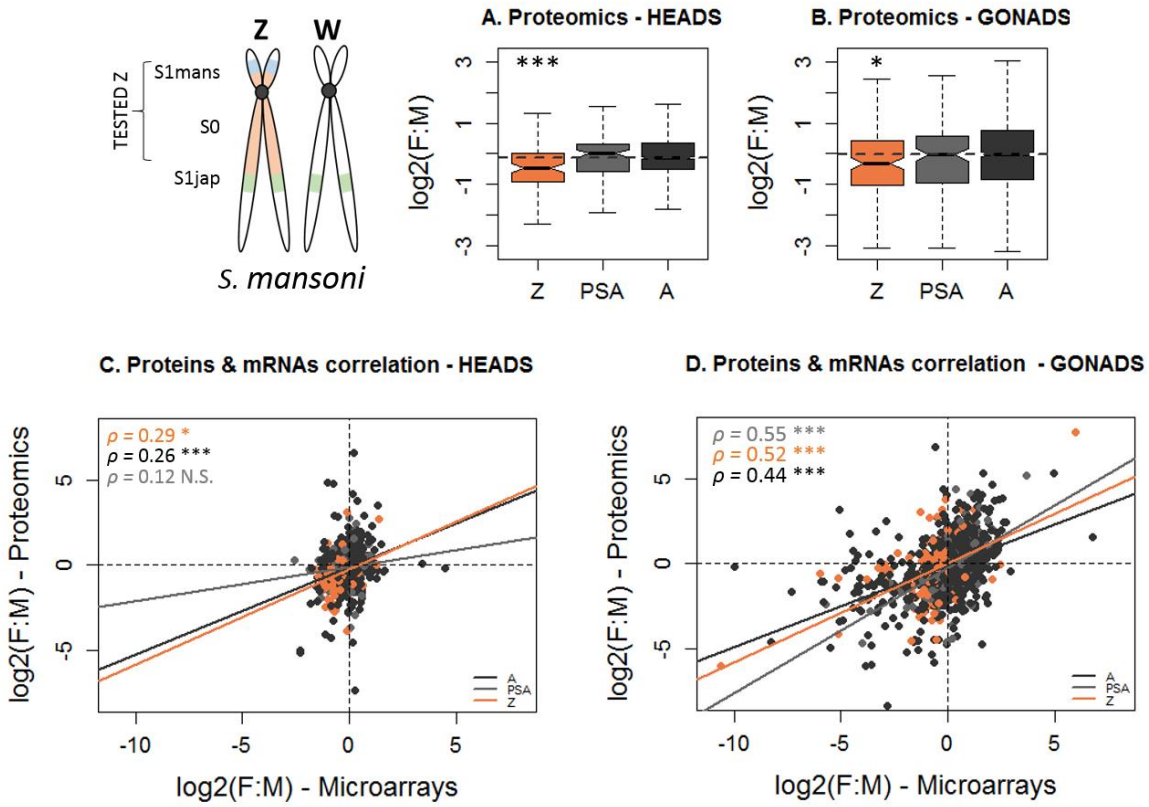


Figure 4. Transcript and protein dosage of Z-linked and autosomal genes in *S. mansoni* heads and gonads. (A), (B) The female:male (F:M) ratio of protein dosage in *S. mansoni* heads **(A)** and gonads **(B)**, for Z-specific (Z), pseudoautosomal (PSA) or autosomal (A) genes. The dotted line shows the autosomal median of F:M expression. The level of significance of each comparison (Wilcoxon rank sum test with continuity correction) is denoted with asterisks: **P-value*<0.05, ***P-value*<0.001, ****P-value*<0.0001. **(C), (D)** Pearson correlation between the female:male ratio of expression obtained by proteomics (y-axes) and by microarrays (x-axes) in *S.mansoni* heads **(C)** and gonads **(D).** A positive correlation (coefficients ρ) is observed for Z-linked (Z, in orange), pseudoautosomal (PSA, in grey) and autosomal genes (A, in darkgrey). The level of significance of each correlation is denoted by asterisks: **P-value*<0.05, ***P-value*<0.001, ****P-value*<0.0001, N.S. *P-value*>0.05. No significant difference was found between the correlation obtained for Z-linked and autosomal genes in either tissue (using a Fisher r-to-z transformation).

Legends of Appendix 1 Figures and Supplementary Figures

Appendix 1 - Figure legends

Appendix 1 - Figure 1. Determination of Z-specific maximum coverage threshold for S. mansoni reference species

Appendix 1 - Figure 2. Determination of Z-specific maximum coverage threshold for S. japonicum species

Appendix 1 - Figure 3. Determination of Z-specific maximum coverage threshold for S. haematobium species

Appendix 1 - Figure 4. Correlation analysis between RNAseq libraries: heatmaps

Supplementary Figures

Figure 2-figure supplement 1: Adult expression patterns (in RPKM) of genes located in the different strata of the Z (S0, S1man and S1jap), as well as pseudoautosomal and autosomal genes. The distribution of expression values is shown for (A) S. mansoni females, (B) S. mansoni males, (C) S. mansoni female:male, (D) S. japonicum females, (E) S. japonicum males, (F) S. japonicum female:male, (G) Female expression in S. japonicum divided by female expression in S. mansoni, (H) Male expression in S. japonicum divided by male expression in S. mansoni, (I) female:male ratio of expression in S. japonicum divided by female:male ratio of expression in S. mansoni. Asterisks represent levels of significance of comparisons to the control autosomal genes (*P-value<0.05, **P-value<0.001, ***P-value<0.0001). Genes were assigned to each Z-chromosome stratum based only on their coverage patterns; only genes with RPKM>1 were used.

Figure 2-figure supplement 2: Adult expression patterns (in RPKM) of genes located in the different strata of the Z (S0, S1man and S1jap), as well as pseudoautosomal and autosomal genes. The distribution of expression values is show for (A) S. mansoni females, (B) S. mansoni males, (C) S. mansoni female:male, (D) S. japonicum females, (E) S. japonicum males, (F) S. japonicum female:male, (G) Female expression in S. japonicum divided by female expression in S. mansoni, (H) Male expression in S. japonicum divided by male expression in S. mansoni, (I) female:male ratio of expression in S. japonicum divided by female:male ratio of expression in S. mansoni. Asterisks represent levels of significance of comparisons to the control autosomal genes (*P-value<0.05, **P-value<0.001, ***P-value<0.0001). Genes were assigned to each Z-chromosome stratum based only on their coverage patterns; only genes with RPKM>0 were used.

Figure 2-figure supplement 3: Adult expression patterns (in RPKM) of genes located in the different strata of the Z (S0, S1man and S1jap), as well as pseudoautosomal and autosomal genes. The distribution of expression values is show for (A) S. mansoni females, (B) S. mansoni males, (C) S. mansoni female:male, (D) S. japonicum females, (E) S. japonicum males, (F) S. japonicum female:male, (G) Female expression in S. japonicum divided by female expression in

S. mansoni, (H) Male expression in S. japonicum divided by male expression in S. mansoni, (I) female:male ratio of expression in S. japonicum divided by female:male ratio of expression in S. mansoni. Asterisks represent levels of significance of comparisons to the control autosomal genes (*P-value<0.05, **P-value<0.001, ***P-value<0.0001). Genes previously mapped to the ZW linkage group were assigned to each Z-chromosome stratum based on their coverage patterns (stringent classification); only genes with RPKM>1 were used.

Figure 2-figure supplement 4: Adult expression patterns (in RPKM) of genes located in the different strata of the Z (S0, S1man and S1jap), as well as pseudoautosomal and autosomal genes. The distribution of expression values is show for (A) S. mansoni females, (B) S. mansoni males, (C) S. mansoni female:male, (D) S. japonicum females, (E) S. japonicum males, (F) S. japonicum female:male, (G) Female expression in S. japonicum divided by female expression in S. mansoni, (H) Male expression in S. japonicum divided by male expression in S. mansoni, (I) female:male ratio of expression in S. japonicum divided by female:male ratio of expression in S. mansoni. Asterisks represent levels of significance of comparisons to the control autosomal genes (*P-value<0.05, **P-value<0.001, ***P-value<0.0001). Genes previously mapped to the ZW linkage group were assigned to each Z-chromosome stratum based on their coverage patterns (stringent classification); only genes with RPKM>0 were used.

Figure 2-figure supplement 5: Adult expression patterns (in RPKM) of genes located in the different strata of the Z (S0, S1man and S1jap), as well as pseudoautosomal and autosomal genes. The distribution of expression values is show for (A) S. mansoni females, (B) S. mansoni males, (C) S. mansoni female:male, (D) S. japonicum females, (E) S. japonicum males, (F) S. japonicum female:male, (G) Female expression in S. japonicum divided by female expression in S. mansoni, (H) Male expression in S. japonicum divided by male expression in S. mansoni, (I) female:male ratio of expression in S. japonicum divided by female:male ratio of expression in S. mansoni. Asterisks represent levels of significance of comparisons to the control autosomal genes (*P-value<0.05, **P-value<0.001, ***P-value<0.0001). Genes were assigned to each Z-chromosome stratum based only on their coverage patterns; only genes with RPKM>0 were used. This figure differs from Supplementary Figure 2 as the Wormbase Parasite list 1:1 orthologs was used instead of our own reciprocal best hit results.

Figure 2-figure supplement 6: Adult expression patterns (in TPM) of genes located in the different strata of the Z (S0, S1man and S1jap), as well as pseudoautosomal and autosomal genes. The distribution of expression values is show for (A) S. mansoni females, (B) S. mansoni males, (C) S. mansoni female:male, (D) S. japonicum females, (E) S. japonicum males, (F) S. japonicum female:male, (G) Female expression in S. japonicum divided by female expression in S. mansoni, (H) Male expression in S. japonicum divided by male expression in S. mansoni, (I) female:male ratio of expression in S. japonicum divided by female:male ratio of expression in S. mansoni. Asterisks represent levels of significance of comparisons to the control autosomal genes (*P-value<0.05, **P-value<0.001, ***P-value<0.0001). Genes were assigned to each Z-chromosome stratum based only on their coverage patterns; only genes with TPM>1 were used.

Figure 2-figure supplement 7: Adult expression patterns (in TPM) of genes located in the different strata of the Z (S0, S1man and S1jap), as well as pseudoautosomal and autosomal genes. The distribution of expression values is show for (A) S. mansoni females, (B) S. mansoni males, (C) S. mansoni female:male, (D) S. japonicum females, (E) S. japonicum males, (F) S. japonicum female:male, (G) Female expression in S. japonicum divided by female expression in S. mansoni, (H) Male expression in S. japonicum divided by male expression in S. mansoni, (I) female:male ratio of expression in S. japonicum divided by female:male ratio of expression in S. mansoni. Asterisks represent levels of significance of comparisons to the control autosomal genes (*P-value<0.05, **P-value<0.001, ***P-value<0.0001). Genes were assigned to each Z-chromosome stratum based only on their coverage patterns; only genes with TPM>0 were used.

Figure 2-figure supplement 8: Schistosomula expression patterns (in RPKM) of genes located in the different strata of the Z (S0, S1man and S1jap), as well as pseudoautosomal and autosomal genes. The distribution of expression values is show for (A) S. mansoni females, (B) S. mansoni males, (C) S. mansoni female:male, (D) S. japonicum females, (E) S. japonicum males, (F) S. japonicum female:male, (G) Female expression in S. japonicum divided by female expression in S. mansoni, (H) Male expression in S. japonicum divided by male expression in S. mansoni, (I) female:male ratio of expression in S. japonicum divided by female:male ratio of expression in S. mansoni. Asterisks represent levels of significance of comparisons to the control autosomal genes (*P-value<0.05, **P-value<0.001, ***P-value<0.0001). Genes were assigned to each Z-chromosome stratum based only on their coverage patterns; only genes with RPKM>1 were used.

Figure 2-figure supplement 9: Schistosomula expression patterns (in RPKM) of genes located in the different strata of the Z (S0, S1man and S1jap), as well as pseudoautosomal and autosomal genes. The distribution of expression values is show for (A) S. mansoni females, (B) S. mansoni males, (C) S. mansoni female:male, (D) S. japonicum females, (E) S. japonicum males, (F) S. japonicum female:male, (G) Female expression in S. japonicum divided by female expression in S. mansoni, (H) Male expression in S. japonicum divided by male expression in S. mansoni, (I) female:male ratio of expression in S. japonicum divided by female:male ratio of expression in S. mansoni. Asterisks represent levels of significance of comparisons to the control autosomal genes (*P-value<0.05, **P-value<0.001, ***P-value<0.0001). Genes were assigned to each Z-chromosome stratum based only on their coverage patterns; only genes with RPKM>0 were used.

Figure 2-figure supplement 10: Schistosomula expression patterns (in RPKM) of genes located in the different strata of the Z (S0, S1man and S1jap), as well as pseudoautosomal and autosomal genes. The distribution of expression values is show for (A) S. mansoni females, (B) S. mansoni males, (C) S. mansoni female:male, (D) S. japonicum females, (E) S. japonicum males, (F) S. japonicum female:male, (G) Female expression in S. japonicum divided by female expression in S. mansoni, (H) Male expression in S. japonicum divided by male expression in S. mansoni, (I) female:male ratio of expression in S. japonicum divided by female:male ratio of expression in S. mansoni. Asterisks represent levels of significance of comparisons to the

control autosomal genes (*P-value<0.05, **P-value<0.001, ***P-value<0.0001). Genes previously mapped to the ZW linkage group were assigned to each Z-chromosome stratum based on their coverage patterns (stringent classification); only genes with RPKM>1 were used.

Figure 2-figure supplement 11: Schistosomula expression patterns (in RPKM) of genes located in the different strata of the Z (S0, S1man and S1jap), as well as pseudoautosomal and autosomal genes. The distribution of expression values is show for (A) S. mansoni females, (B) S. mansoni males, (C) S. mansoni female:male, (D) S. japonicum females, (E) S. japonicum males, (F) S. japonicum female:male, (G) Female expression in S. japonicum divided by female expression in S. mansoni, (H) Male expression in S. japonicum divided by male expression in S. mansoni, (I) female:male ratio of expression in S. japonicum divided by female:male ratio of expression in S. mansoni. Asterisks represent levels of significance of comparisons to the control autosomal genes (*P-value<0.05, **P-value<0.001, ***P-value<0.0001). Genes previously mapped to the ZW linkage group were assigned to each Z-chromosome stratum based on their coverage patterns (stringent classification); only genes with RPKM>0 were used.

Figure 2-figure supplement 12: Schistosomula expression patterns (in RPKM) of genes located in the different strata of the Z (S0, S1man and S1jap), as well as pseudoautosomal and autosomal genes. The distribution of expression values is show for (A) S. mansoni females, (B) S. mansoni males, (C) S. mansoni female:male, (D) S. japonicum females, (E) S. japonicum males, (F) S. japonicum female:male, (G) Female expression in S. japonicum divided by female expression in S. mansoni, (H) Male expression in S. japonicum divided by male expression in S. mansoni, (I) female:male ratio of expression in S. japonicum divided by female:male ratio of expression in S. mansoni. Asterisks represent levels of significance of comparisons to the control autosomal genes (*P-value<0.05, **P-value<0.001, ***P-value<0.0001). Genes were assigned to each Z-chromosome stratum based only on their coverage patterns; only genes with RPKM>0 were used. This figure differs from Supplementary Figure 2 as the Wormbase Parasite list 1:1 orthologs was used instead of our own reciprocal best hit results.

Figure 2-figure supplement 13: Schistosomula expression patterns (in TPM) of genes located in the different strata of the Z (S0, S1man and S1jap), as well as pseudoautosomal and autosomal genes. The distribution of expression values is show for (A) S. mansoni females, (B) S. mansoni males, (C) S. mansoni female:male, (D) S. japonicum females, (E) S. japonicum males, (F) S. japonicum female:male, (G) Female expression in S. japonicum divided by female expression in S. mansoni, (H) Male expression in S. japonicum divided by male expression in S. mansoni, (I) female:male ratio of expression in S. japonicum divided by female:male ratio of expression in S. mansoni. Asterisks represent levels of significance of comparisons to the control autosomal genes (*P-value<0.05, **P-value<0.001, ***P-value<0.0001). Genes were assigned to each Z-chromosome stratum based only on their coverage patterns; only genes with TPM>1 were used.

Figure 2-figure supplement 14: Schistosomula expression patterns (in TPM) of genes located in the different strata of the Z (S0, S1man and S1jap), as well as pseudoautosomal and autosomal genes. The distribution of expression values is show for (A) S. mansoni females, (B)

S. mansoni males, (C) S. mansoni female:male, (D) S. japonicum females, (E) S. japonicum males, (F) S. japonicum female:male, (G) Female expression in S. japonicum divided by female expression in S. mansoni, (H) Male expression in S. japonicum divided by male expression in S. mansoni, (I) female:male ratio of expression in S. japonicum divided by female:male ratio of expression in S. mansoni. Asterisks represent levels of significance of comparisons to the control autosomal genes (*P-value<0.05, **P-value<0.001, ***P-value<0.0001). Genes were assigned to each Z-chromosome stratum based only on their coverage patterns; only genes with TPM>0 were used.

Figure 2-figure supplement 15. Z-linked and autosomal female:male ratio of gene expression using different filters. The Z-linked and autosomal gene expression patterns are shown in S. japonicum and S. mansoni. Different filters were applied: a minimum threshold of expression of 1 RPKM in all libraries (A-D), a minimum threshold of expression of 3 RPKM in all libraries (E-H), or a minimum threshold of expression of 1 RPKM per stage (as in the main manuscript), together with a filter on fold change to account for strong sex-biases (FC<2 for female:male ratio and male:female ratio) (I¬-L). Z refers to the female:male ratio of expression in the Z-specific region and A refers to the female:male ratio of expression on autosomes. The level of significance of each comparison is denoted with asterisks: ***P-value<0.0001.

Figure 2-figure supplement 16. Z-linked and autosomal gene expression in females and males using different filters. The Z-linked and autosomal gene expression patterns are shown for S. japonicum and S. mansoni. Different filters were applied: a minimum threshold of expression of 1 RPKM in all libraries (A-D), a minimum threshold of expression of 3 RPKM in all libraries (E-H), or a minimum threshold of expression of 1 RPKM per stage (as in the main manuscript), together with a filter on fold change to account for strong sex biases (FC<2 for female:male ratio and male:female ratio) (I¬-L). Fem-Z and Male-Z refers to the female:male ratio of expression in the Z-specific region and Fem-A and Male-A refers to the female:male ratio of expression on autosomes, in female and male respectively The level of significance of each comparison is denoted with asterisks: *P-value<0.05, **P-value<0.001, ***P-value<0.0001.

Figure 2-figure supplement 17. Z-linked and autosomal female:male ratio of gene expression according to presence/absence of known protein-protein interactions. The Z-linked and autosomal gene expression patterns are shown for schistosomula and adults, in S. japonicum (A-B) and S. mansoni (C-D). We considered only genes with minimum level of expression RPKM>1 in all stages, did not take into account the ancestral state of expression, and used the “exhaustive strata” set of genes. -PPIs and +PPIs refer to the female:male ratio of expression in the lineage specific Z-linked strata S1jap and S1mans respectively. The level of significance of each comparison is denoted with asterisks: ***P-value<0.0001.

Figure 3-figure supplement 1. Z-linked and autosomal female:male ratio of gene expression, normalized by ancestral autosomal expression, and using different filters. The Z-linked and autosomal gene expression patterns (normalized by ancestral pseudo-autosomal or autosomal expression, to show changes since the appearance of the two S1 strata) are shown for S1jap in S. japonicum and S1mans in S. mansoni. Different filters were applied: a minimum threshold of

expression of 1 RPKM in all libraries (A-D), a minimum threshold of expression of 3 RPKM in all libraries (E-H), or a minimum threshold of expression of 1 RPKM per stage (as in the main manuscript), together with a filter on fold change to account for strong sex biases (FC<2 for female:male ratio and male:female ratio) (I¬-L). S1jap and S1mans refer to the female:male ratio of expression in the lineage specific Z-linked strata S1jap and S1mans respectively. The level of significance of each comparison is denoted with asterisks: ***P-value<0.0001.

Figure 3-figure supplement 2. Z-linked and autosomal gene expression in females and males, normalized by ancestral autosomal expression, and using different filters. The Z-linked and autosomal gene expression patterns (normalized by ancestral pseudo-autosomal or autosomal expression, to show changes since the appearance of the two S1 strata) are shown for S1jap in S. japonicum and S1mans in S. mansoni. Different filters were applied: a minimum threshold of expression of 1 RPKM in all libraries (A-D), a minimum threshold of expression of 3 RPKM in all libraries (E-H), or a minimum threshold of expression of 1 RPKM per stage (as in the main manuscript), together with a filter on fold change to account for strong sex biases (FC<2 for female:male ratio and male:female ratio) (I¬-L). Fem-S1j and Male-S1j refer to the normalized expression levels of genes in the stratum S1jap in females and males, Fem-S1m and Male-S1m refer to the normalized expression levels in S1mans, and Fem-A and Male-A refer to the normalized expression levels of autosomal genes in females and in males, respectively. The level of significance of each comparison is denoted with asterisks: *P-value<0.05, **P-value<0.001, ***P-value<0.0001.

Figure 4-figure supplement 1: Pearson correlations between gene dosage at the transcript and protein levels in male heads and gonads. A positive correlation (coefficients R) is observed for Z-linked (Z, in light blue), pseudo-autosomal (PSA, in grey) and autosomal genes (A, in dark blue). The level of significance of each correlation is denoted by asterisks: *P-value<0.05, **P-value<0.001, ***P-value<0.0001. No significant difference was found between the correlations obtained for Z-linked and autosomal genes in either tissue (using a Fisher r-to-z transformation).

Figure 4-figure supplement 2: Pearson correlations between gene dosage at the transcript and protein levels in female heads and gonads. A positive correlation (coefficients R) is observed for Z-linked (Z, in light pink), pseudo-autosomal (PSA, in grey) and autosomal genes (A, in dark pink. The level of significance of each correlation is denoted by asterisks: *P-value<0.05, **P-value<0.001, ***P-value<0.0001. No significant difference was found between the correlations obtained for Z-linked and autosomal genes in either tissue (using a Fisher r-to-z transformation).

Figure 4-figure supplement 3: Pearson correlations between the female:male ratio of expression obtained by proteomics (y-axes) and by microarrays (x-axes) in S.mansoni heads and gonads, using only genes with male-biased expression in the microarray data. A positive correlation (coefficients R) is observed for Z-linked (Z, in orange), pseudo-autosomal (PSA, in grey) and autosomal genes (A, in darkgrey). The level of significance of each correlation is denoted by asterisks: *P-value<0.05, **P-value<0.001, ***P-value<0.0001, N.S. P-value>0.05.

No significant difference was found between the correlations obtained for Z-linked and autosomal genes in either tissue (using a Fisher r-to-z transformation).

Source data files

Figure 1-source data 1: Comparative genomics: coverage analysis and strata identification

Figure 1-source data 2: S. mansoni reference species: coverage analysis and Z-Autosome assignment

Figure 1-source data 3: S. haematobium autosome vs Z-specific region assigment

Figure 1-source data 4: S. japonicum autosome vs Z-specific region assignment

Figure 2-source data 1: RPKM calculation for S. mansoni gene expression

Figure 2-source data 2: RPKM calculation for S. japonicum gene expression

Figure 3-source data 1: One-to-one orthology S. mansoni vs S. japonicum

Figure 3-source data 2: Transcriptomic data, for blat 1-to-1 orthologs

Figure 3-source data 3: Transcriptomic data, for WormBase Biomart orthologs

Figure 4-source data 1: Proteomic data with imputed values – GONADS

Figure 4-source data 2: Proteomic data with imputed values – HEADS

Figure 4-source data 3: Proteomic data without imputed values – GONADS

Figure 4-source data 4: Proteomic data without imputed values – HEADS

Figure 4-source data 5: Correponsdance between Gene_Id and Protein_ID

Figure 4-source data 6: Microarray data

Supplementary File 1. Libraries and captions for source data files.

Appendix 1 – Computational analyses

Supplementary codes and inputs

Additional Perl and R scripts are available at the following link:
https://www.dropbox.com/sh/1jpdp15zokx7z85/AAB4BoRm5Jf6i7qEys8x_ERMa?dl=0
Remarks: Deposited on IST Austria data repository (http://dx.doi.org/10.15479/AT:ISTA:109)

Additional input files are available at the following link:
https://www.dropbox.com/sh/r0gz93xbkgcsxyq/AABJqAarTNWewqH-Nd_LF1y-a?dl=0
Remarks: Deposited on IST Austria data repository (http://dx.doi.org/10.15479/AT:ISTA:109)

Publicly available reference genomes and annotations

All data were obtained at the WormBase Parasite database (https://parasite.wormbase.org/index.html)(56,57).
S. mansoni:
Genomic sequences: schistosoma_mansoni.PRJEA36577.WBPS9.genomic.fa
Annotation: schistosoma_mansoni.PRJEA36577.WBPS9.canonical_geneset.gtf
CDS: schistosoma_mansoni.PRJEA36577.WBPS9.CDS_transcripts.fa
S. japonicum:
Genomic sequences: schistosoma_japonicum.PRJEA34885.WBPS9.genomic.fa
Annotation: schistosoma_japonicum.PRJEA34885.WBPS9.canonical_geneset.gtf
CDS: schistosoma_japonicum.PRJEA34885.WBPS9.CDS_transcripts.fa
S. haematobium:
Genomic sequences: schistosoma_haematobium.PRJNA78265.WBPS9.genomic.fa
Annotation: schistosoma_haematobium.PRJNA78265.WBPS9.canonical_geneset.gtf

Orthology

The same pipeline was applied to *S. japonicum* and *S. haematobium*. Example is shown here for the *S. japonicum* species.

Mapping of *S. mansoni* genes on the *S. japonicum* genome
First, *S. mansoni* gene sequences were mapped to the *S. japonicum* genome using Blat (79) with a translated query and database:

```
CDS=pathwaytosmansonigenome/schistosoma_mansoni.PRJEA36577.WBPS9.CDS_transcripts.fa
GENOME=~pathwaytojaponicumgenome
BLATOUTPUT=~pathwaytoblatoutputfile

~/tools/blat -q=dnax -t=dnax -minScore=50
${GENOME}/schistosoma_japonicum.PRJEA34885.WBPS9.genomic.fa ${CDS}
${BLATOUTPUT}/jap_mans.blat
```

The resulting blat alignment was filtered to keep only the mapping hit with the highest mapping for each *S. mansoni* gene:

```
sort -k 10 ${BLATOUTPUT}/jap_mans.blat > ${BLATOUTPUT}/jap_mans.blat.sorted

perl Script1_besthitblat.pl ${BLATOUTPUT}/jap_mans.blat.sorted
```

This produced the first filtered Blat alignment:

```
${BLATOUTPUT}/jap_mans.blat.sorted.besthit
```

In a second filtering step, when several *S. mansoni* genes overlapped on the *S. japonicum* genome by more than 20bps, we keep only the gene with the highest mapping score:

```
sort -k 14 ${BLATOUTPUT}/jap_mans.blat.sorted.besthit >
${BLATOUTPUT}/jap_mans.blat.sorted.besthit.sorted

perl Script2_blatreverse.pl ${BLATOUTPUT}/jap_mans.blat.sorted.besthit.sorted
```

This produced the final filtered Blat alignment:

```
${BLATOUTPUT}/jap_mans.blat.sorted.besthit.sorted.nonredundant
```

Assignment of syntenic blocks to *S. mansoni* chromosomal locations

Each *S. japonicum* scaffold was assigned to one of the *S. mansoni* chromosomes based on its gene content, using a majority rule.

(i) As input table, we transformed the output obtained in the previous step into a 3-column file (sorted by column 1): column1=Smansoni_gene_ID, column2=Sjaponicum_contig/scaffold, column3=score.

```
cat ${BLATOUTPUT}/jap_mans.blat.sorted.besthit.sorted.nonredundant | cut -f 1,10,14 |
awk '{print $2, $3, $1}' | perl -pi -e 's/-P.//gi' | sort >
${BLATOUTPUT}/jap_mans.blat.sorted.besthit.sorted.nonredundant_score.joignabl
```

(ii) A second file containing the chromosomal location of *S. mansoni* genes was also needed for the next step. The file (named here RefChromLocation.joignabl) is a 2-column file resulting from the combination of the *S. mansoni* GFF file, and the newly identified Z-specific genes (*i.e.* the chromosomal location corresponds to the published GTF location, except for 1. the new Z *versus* pseudoautosomal assignments based on coverage for ZW genes and 2. genes which were newly detected as Z-linked): column1=Smansoni_gene_ID, column2=Smansoni_chromosome. This file should be sorted by column 1. Column2 has 11 possible values: "Chr_1", "Chr_2", "Chr_3","Chr_4", "Chr_5", "Chr_6", "Chr_7", "Chr_PSA", "Chr_Z", "Chr_newZ", and "Chr_UP" (See Figure 1-source data 1, column #6: “Bestloc_input_mans”).

(iii) We joined both files by column1 and sort the joined table by Sjaponicum_contig/scaffold:

```
join ${BLATOUTPUT}/jap_mans.blat.sorted.besthit_score.joignabl pathwaytosmansonigenome/
RefChromLocation.joignabl | awk '{print $2, $1, $4, $3}' | sort >
jap_mans.inputforbestlocation
```

A final script counted how many orthologs from the different *S. mansoni* chromosomes were found in each *S. japonicum* scaffold, and assigned it a chromosome based on the location of the majority of the orthologs. If the same number of genes were located on two different chromosomes, then the chromosome for which the sum of the mapping scores was highest waschosen.

```
perl Script3_bestlocation.pl jap_mans.inputforbestlocation
```

Identification of 1-to-1 orthologs between *S. mansoni* and *S. japonicum*

S. japonicum genes were first mapped to *S. mansoni* using Blat with a translated database and query:

```
CDS1=pathwaytosmansonigenome/schistosoma_mansoni.PRJEA36577.WBPS9.CDS_transcripts.fa
CDS2=pathwaytojaponicumgenome/schistosoma_japonicum.PRJEA34885.WBPS9.CDS_transcripts.fa
BLATOUTPUT=~pathwaytoblatoutputfile

~/tools/blat -q=dnax -t=dnax -minScore=50 ${CDS2} ${CDS1}
${BLATOUTPUT}/jap_mans_onetoone.blat
```

Only best reciprocal hits were kept:

```
perl Script4_Blat_one2one.pl ${BLATOUTPUT}/jap_mans_onetoone.blat
```

The list of orthologs is provided in Figure 3-source data 1.

Determination of Z-specific regions in each species

Raw read processing

The same pipeline was applied to all the species. Example is shown here for the *S. japonicum* species.
The newly sequenced *S. japonicum* reads were provided as BAM files, which we first converted to FASTQ:

```
READS=~pathwaytoreads

for i in `ls ${READS} | grep .bam`; do
echo ${i}
```

```
~/tools /bedtools.2.25.0/bin/bamToFastq -i ${i} -fq ${i}.fastq1 -fq2 ${i}.fastq2
done
```

Adapter sequences were removed from the newly sequenced libraries (Only for *S. japonicum*) using Cutadapt (80):

```
LIBRARY_1=Forward reads in fastq format
LIBRARY_2=Reverse reads in fastq format
READS =~pathwaytoreads
CUTADAPTOUTPUT=~pathwaytocutadaptoutput

~/tools/cutadapt_1.9.1 --match-read-wildcards -f fastq -a
AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC ${READS}/${LIBRARY_1} -o
${CUTADAPTOUTPUT}/${LIBRARY_1}.trimmed

~/tools/cutadapt_1.9.1 --match-read-wildcards -f fastq -a
AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTA ${READS}/${LIBRARY_2} -o
${CUTADAPTOUTPUT}/${LIBRARY_2}.trimmed
```

Trimmomatic was used to remove adaptor and low quality sequences from the available libraries (For *S. mansoni* and *S. haematobium*, accession numbers are given in the main text):

```
READS=~pathwaytoreads
LIBRARY=LibraryName

module load java
java -jar ~/tools/tools/Trimmomatic-0.36/trimmomatic-0.36.jar PE -phred33
${READS}/${LIBRARY}.fastq1 ${READS}/${LIBRARY}.fastq2
${READS}/${LIBRARY}.fastq1_paired.fq.gz ${READS}/${LIBRARY}.fastq1_unpaired.fq.gz
${READS}/${LIBRARY}.fastq2_paired.fq.gz ${READS}/${LIBRARY}.fastq2_unpaired.fq.gz
ILLUMINACLIP:~/tools/Trimmomatic-0.36/adapters/TruSeq2-PE.fa:2:30:10 HEADCROP:12
LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36
```

All libraries were checked using FastQC:

```
for i in `ls ${CUTADAPTOUTPUT} | grep .trimmed`; do
~/tools/FastQC-v0.11.2/fastqc ${CUTADAPTOUTPUT}/${i} --extract –o ~/fastQCoutputs/
done
```

OR

```
for i in `ls ${READS} | grep paired.fq.gz`; do
~/tools/FastQC-v0.11.2/fastqc ${READS}/${i} --extract –o ~/fastQCoutputs/
done
```

Read alignment on reference genome

The same pipeline was applied to all the species. Example is shown here for the *S. japonicum* species.
Reads were mapped to the respective reference genomes using Bowtie2 (82).

```
#Built reference genome

GENOME=~pathwaytojaponicumgenome

~/tools/bowtie2-2.2.9/bowtie2-build
${GENOME}/schistosoma_japonicum.PRJEA34885.WBPS9.genomic.fa ref_genome

# Mapping

SPE=Sjaponicum
LIBRARY=LibraryName
BOWTIE2OUTPUT=~pathwaytobowtie2output
THREADS=5

~/tools/bowtie2-2.2.9/bowtie2 -x ${GENOME}/refgenome -1
${CUTADAPTOUTPUT}/${LIBRARY_1}.trimmed -2 ${CUTADAPTOUTPUT}/${LIBRARY_2}.trimmed --end-
to-end --sensitive -p ${THREADS} -S ${BOWTIE2OUTPUT}/${SPE}_${LIBRARY}.sam
```

Genomic coverage

The same pipeline was applied to all the species. Example is shown here for the *S. japonicum* species.
Genomic coverage in the male and female samples was calculated using SOAPcoverage (v2.7.7., http://soap.genomics.org.cn/index.html), after filtering for uniquely mapped reads.

```
#Select only uniquely mapped reads

cd ${BOWTIE2OUTPUT}

for i in `ls | grep .sam`; do
echo ${i}
grep -vw "XS:i" ${i} > ${i}_unique.sam
done

#Obtain average genomic coverage for each scaffold

GENOME=~pathwaytojaponicumgenome
SOAPCOVOUTPUT=pathwaytosoapcoverageoutput

cd ${BOWTIE2OUTPUT}

for j in `ls | grep _unique.sam`; do
echo ${j}
~/tools/soapcoverage.2.7.7/soap.coverage -sam -cvg -i ${j} -onlyuniq -refsingle
${GENOME} -o ${SOAPCOVOUTPUT}/${j}.soapcov
done
```

For *S. mansoni,* the genomic coverage was not calculated per scaffold but for 10kb windows:

```
GENOME=~pathwaytomansonigenome
LIBRARY=LibraryName
SOAPCOVOUTPUT=pathwaytosoapcoverageoutput
WIN=10000

cd ${BOWTIE2OUTPUT}

for j in `ls | grep _unique.sam`; do
echo ${j}
~/tools/soapcoverage.2.7.7/soap.coverage -sam -cvg -i ${j} -onlyuniq -refsingle
${GENOME} -window ${WIN} -o ${SOAPCOVOUTPUT}/${j}.soapcov
done
```

The final coverage values are part of the Figure 1-source data 2 (*S. mansoni*), Figure 1-source data 3 (*S. haematobium*) and Figure 1-source data 4 (*S. japonicum*)

Determination of Z-specific maximum coverage threshold for *S. mansoni* reference species

The maximum value of log2(female:male) coverage for Z-specific assignment in *S. mansoni* was determined using the R script Script5_Genomics_mansoni (Part I) with the input file InputFile1_Smansoni_SoapCov_By10kb_ForR, which contains the SoapCov outputs for male (ERR562989_Depth) and female (ERR562990_Depth).

This involved the following steps (shown in Appendix 1-figure 1):

A) When plotted, the log2(female:male) ratio of coverage shows a bimodal distribution. The highest peak is distributed around 0 and corresponds to the windows that are localized on the known autosomes (Chromosomes 1 to 7) of the current version of the genome assembly (Protasio et al. 2012 (42), PRJEA36577).
B) If we consider only the unimodal distribution of autosomal window coverage, the 5th percentile has a log2(female:male)=-0.26.
C) We use this value to exclude the pseudoautosomal windows (plotted in grey) of the ZW linkage group and consider only the non-pseudoautosomal windows (plotted in orange) for the next step.
D) The non-pseudoautosomal window (InputFile2_Z_10kb_5pr100) coverage displays an unimodal distribution, and the 99th percentile as a value log2(female:male)=-0.4.
E) When visualizing log2(female:male) along the ZW linkage group, it appears that the log2(female:male)=-0.4 threshold allows the discrimination of the three known Z-specific regions (in orange) and pseudoautosomal regions (in grey). To finely define the Z-specific content, we systematically excluded windows for which the

coverage value is not consistent with the adjacent windows: 5 consecutive 10kb windows with consistent coverage were required to be considered either pseudoautosomal or Z-specific. For instance, yellow bands on the graph highlight regions of more than one window but less than five. They are tagged as "Excluded" in Figure 1-source data 2, as are the two adjacent windows.

We then applied this threshold value to all the genome, considering only scaffolds longer than 50kb for classification. The final classification of each scaffold (as "Z" or "Autosome") is presented in Figure 1-source data 2. Used R script is: Script1_Genomics_mansoni (Part II); and corresponding input file is: InputFile1_Smansoni_SoapCov_By10kb_ForR.
All genes of a window were assigned to "Z" or "Autosome" depending on the window assignment. When a gene overlapped with an excluded window, it was excluded as well. In order to exclude these genes, the following script was used: Script6_GeneSelector, using as input a 3-column file (Chromosome ID, First base of window to exclude, Last base of window to exclude).

Determination of Z-specific maximum coverage threshold for *S. japonicum.*
The maximum value of log2(female:male) coverage for Z-specific assignment in *S. japonicum* was determined using the R script Script7_Genomics_japonicum (Part I) with the input file InputFile3_Sjaponicum_SoapCov_Bestloc_ForR, which contains the SoapCov outputs for male (Depth_40640) and female (ERR562990_Depth), and the assignment to *S. mansoni* chromosomes (See paragraph 1.1).

This involved the following steps (shown in Appendix 1-figure 2):

A) When plotted, the log2(female:male) ratio of coverage shows a bimodal distribution.
B) As *S. japonicum* genome is only assembled at the scaffold level (Zhou et al. 2009), we used the log2(female:male) coverage of the scaffold mapping to *S. mansoni* autosomes to get the distribution of autosomal log2(female:male).
C) If we consider only the unimodal distribution of the autosome-assigned scaffold coverage, the 1st percentile has a log2(female:male)=-0.4.
D) We use this value to classify scaffolds that mapped to the *S. mansoni* ZW linkage group as Z-specific or pseudoautosomal (grey in Appendix 1-figure 2, panel D), and consider only the non-pseudoautosomal scaffolds (plotted in orange) for the next step.
E) The non-pseudoautosomal scaffold (InputFile4_Sjaponicum_1pr100) coverage displays an unimodal distribution, and the 95th percentile has a value of log2(female:male)=-0.84.
F) When visualizing the log2(female:male) along the ZW linkage group, it appears that no single threshold value can finely discriminate Z-specific and pseudoautosomal scaffolds. So we use the log2(female:male)=-0.84 threshold as maximum value for Z-specific regions (in orange), and log2(female:male)=-0.4 threshold as minimum value for pseudoautosomal regions (in grey). Scaffolds with log2(female:male) between these two values were classified as "ambiguous".

We then applied this threshold value to all the genome. Used R script is : Script7_Genomics_japonicum (Part II); and corresponding input file is: InputFile3-2_Sjaponicum_SoapCov_ForR.
The final "Z", "Autosome" or "Ambiguous" assignment is shown in Figure 1-source data 4.

Determination of Z-specific maximum coverage threshold for *S. haematobium.*
The maximum value of log2(Female) coverage for Z-specific assignment in *S. haematobium* was determined using the R script Script8_Genomics_haematobium (Part I) with the input file InputFile5_Shaematobium_SoapCov_Bestloc_ForR, which contains the SoapCov outputs for two female libraries (Depth_ERR037800, Depth_ERR036251), and the assignment to *S. mansoni* chromosomes (See paragraph 1.1).

This involved the following steps (shown in Appendix 1-figure 3):

A) When plotted, the log2(Female) coverage shows a bimodal distribution.
B) As the *S. haematobium* genome is only assembled at the scaffold level (Young et al. 2012), we used the log2(Female) coverage of the scaffold mapping to *S. mansoni* autosomes to get the distribution of autosomal log2(Female).
C) If we consider only the unimodal distribution of the autosome-assigned scaffold coverage, the 5th percentile has a log2(Female)=4.4
D) We use this value to classify scaffolds that mapped to the *S. mansoni* ZW linkage group as Z-specific or pseudoautosomal (grey in Appendix 1-figure 3, panel D), and consider only the non-pseudoautosomal scaffolds (plotted in orange) for the next step.
E) The non-pseudoautosomal scaffold (InputFile6_Shaematobium_5pr100) coverage displays a unimodal distribution, and the 99th percentile as a value log2(Female)=4.6
G) When visualizing log2(Female) along the ZW linkage group, it appears no single threshold value cannot finely discriminate between Z-specific and pseudoautosomal scaffolds. So we used the log2(Female)=4.6 threshold as maximum value for Z-specific regions (Appendix 1-figure 3, panel F, in orange), and log2(Female)=4.4 threshold as minimum value for pseudoautosomal regions (in grey). Scaffolds with log2(Female) between these two values were classified as "ambiguous".

We then applied this threshold value to all the genome. Used R script is: Script8_Genomics_haematobium (Part II); and corresponding input file is: InputFile5-2_Shaematobium_SoapCov_ForR.txt.
The final "Z", "Autosome" or "Ambiguous" assignment is shown in Figure 1-source data 3.

Definition of evolutionary strata

Genes that were Z-specific in both *S. mansoni* and *S. japonicum* were assigned to stratum S0. Genes that were Z-specific only in *S. mansoni* were assigned to stratum S1mans, and genes that were Z-specific in *S. japonicum* were assigned to stratum S1jap.
This assignment was performed separately for two classifications. The exhaustive classification is defined by the comparison of "Z" or "Autosome" assignment, based on coverage (see 2.4 - 2.6), between *S. mansoni* and *S. japonicum* species. The stringent classification is a subset of the exhaustive classification: it only contains genes localized on the ZW linkage map in *S. mansoni* and with a best location on the Z or PSA region in *S. japonicum* (see Figure 1-source data 1).

Transcriptomics

Raw read processing

Reads were first trimmed with Trimmomatic (v0.36, (81)), using the following commands:

(i) *S. mansoni*

```
READS=~pathwaytoreads

cd ${READS}

for i in `ls | grep .fastq`; do
            echo ${i}
module load java
java -jar ~/tools/Trimmomatic-0.36/trimmomatic-0.36.jar PE -phred33 ${i}
${j}.trimmed.fq.gz ILLUMINACLIP:~/tools/Trimmomatic-0.36/adapters/TruSeq3-SE.fa:2:30:10
HEADCROP:12 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36
done
```

(ii) *S. japonicum*

```
READS=~pathwaytoreads

cd ${READS}

for i in `ls | grep .fastq`; do
            echo ${i}
module load java
java -jar ~/tools/Trimmomatic-0.36/trimmomatic-0.36.jar SE -phred33 ${i}
${j}.trimmed.fq.gz ILLUMINACLIP:~/tools/Trimmomatic-0.36/adapters/TruSeq3-SE.fa:2:30:10
HEADCROP:12 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36
done
```

Read mapping on reference genome

Reads were mapped to their respective reference genomes with TopHat2 (83), an alignment program for RNA-seq reads which takes splice junctions into account.

(i) *S. mansoni*

```
GENOME=~pathwaytogenome/schistosoma_mansoni.PRJEA36577.WBPS9.genomic.fa
READS=~pathwaytoreads
OUF=~/TophatOutput

# Build the indexed genome for Bowtie2

~/tools/bowtie2-2.2.9/bowtie2-build ${genome} refgenome

# Map the RNAseq reads against the reference genome with TopHat2

cd ${READS}
```

```
for i in `ls | grep trimmed.fq`; do
echo ${i}

~tools/tophat-2.1.1.Linux_x86_64/tophat -p 3 --library-type fr-firststrand --microexon-
search -i 10 -I 40000 --min-segment-intron 10 --max-segment-intron 40000 -g 1 -o ${OUF}
~pathwaytogenome/refgenome ${READS}/${i}

done
```

(ii) *S. japonicum*

```
GENOME=~pathwaytogenome/schistosoma_japonicum.PRJEA34885.WBPS9.genomic.fa
READS=~pathwaytoreads
OUF=~/TophatOutput

~tools/tophat-2.1.1.Linux_x86_64/tophat -p 3 --microexon-search -i 10 -I 40000 --min-
segment-intron 10 --max-segment-intron 40000 -g 1 -o ${OUF} ~pathwaytogenome/refgenome
${READS}/${i}
```

Read count

Reads counts for each gene were obtained from the TopHat2 alignments using HTseq (84).

(i) *S. mansoni*

```
INF=~/TopHat_Output
GENOME=~pathwaytogenome/schistosoma_mansoni.PRJEA36577.WBPS9.genomic.fa
OUF=~/HTseq_Output

cd ${INF}

for i in `ls | grep accepted_hits.bam`; do
      echo ${i}
module load anaconda

htseq-count -f bam -s reverse -m union --idattr gene_id -o ${OUF}/${i}_Rev.htseq ${i}
${GENOME} > ${OUF}/${i}_Rev.count
done
```

(ii) *S. japonicum*

```
INF=~/TopHat_Output
GENOME=~/pathwaytogenome/schistosoma_japonicum.PRJEA34885.WBPS9.genomic.fa
OUF=~/HTseq_Output

cd ${INF}

for i in `ls | grep accepted_hits.bam`; do
      echo ${i}
module load anaconda

htseq-count -f bam -s no -m union --idattr gene_id -o ${OUF}/${i}_Rev.htseq ${i}
${GENOME} > ${OUF}/${i}_Rev.count
done
```

Alternative pipeline: Kallisto (example shown for *S. mansoni*)

In order to check that our results held independent of the pipeline used to infer expression levels, we further estimated TPM values using Kallisto (85).

```
# Create index

INF=~pathwaytogenome
CDS=${inf}/schistosoma_mansoni.PRJEA36577.WBPS9.CDS_transcripts.fa # or input for S.
japonicum would be:schistosoma_japonicum.PRJEA34885.WBPS9.CDS_transcripts.fa
OUF=~/Kallistooutput

cd ${OUF}
```

```
~/tools/kallisto_linux-v0.43.1/kallisto index -i transcripts.idx ${CDS}

# Run Kallisto

READS=~pathwaytoreads
OUF=~/Kallistooutput
LIST=Input text file containing the library_ID and located in OUF

cd ${READS}

while read i; do
mkdir ${OUF}/${i}

~/tools/kallisto_linux-v0.43.1/kallisto quant -t 16 -i ${OUF}/transcripts.idx -o
${OUF}/${i} -b 100 --fr-stranded ${READS}/${i}_forward ${READS}/${i}_reverse

done < ${LIST}
```

Comparative analysis of gene expression

We perform all the expression analyses in R (Script9_Transcriptomics) with the input files (i) InputFile7_Transcriptomics1 corresponding to the newly identified one-to-one orthologs (see 1.2); and (ii) InputFile10_Transcriptomics2 corresponding to the list obtained on WormBase parasite.

In adults, a correlation analysis revealed an inconsistency between replicates, corresponding to the *S. mansoni* female and male adults: ERR506076 and ERR506082. The corresponding heatmaps are shown in Appendix 1-figure 4. These two libraries were excluded from further analyses.

Tables resulting from the Loess normalization are InputFile8_SuperTableNormalized_Transcriptomics1_RPKM, InputFile8-2_SuperTableNormalized_LoessOnAll_Transcriptomics1_RPKM, InputFile9_SuperTableNormalized_Transcriptomics1_TPM, and InputFile11_SuperTableNormalized_Transcriptomics2_RPKM.
They can be used directly in the script Script9_Transcriptomics at the step C.

Number of Protein-Protein Interactions (PPIs)

The full set of known PPIs for *S. mansoni* was downloaded from StringDB:
https://stringdb-static.org/download/protein.links.full.v10.5/6183.protein.links.full.v10.5.txt.gz

We kept only interactions supported by experimental evidence or text mining:

```
cat 6183.protein.links.full.v10.5.txt | awk '($10>0 || $14>0)' | awk '{print $1, $2}' |
perl -pi -e 's/6183\.//gi' | perl -pi -e 's/__mRNA//gi' | perl -pi -e 's/\.[0-9]*//gi'
> 6183.protein.links.full.v10.5_exp.txt
```

We cleaned the coverage location, so that "Not analysed (Scaffold <40Kb)" was replaced by NA:

```
cat InputFile14_PPIs_woNA | perl -pi -e 's/Not analysed \(Scaffold \< 40kb\)/NA/gi' |
sort > InputFile14_PPIs
```

Then counted the number of interactions present for each gene:

```
perl Script11_PPIcounter InputFile14_PPIs.txt 6183.protein.links.full.v10.5_exp.txt
```

Output: 5 last columns of "InputFile8-2_SuperTableNormalized_LoessOnAll_Transcriptomics1_RPKM" ($Autocount_exp $Zcount_exp $Excludedcount_exp $NAcount_exp $TotalPPIs_exp)

Proteomics *vs* microarrays correlation

We performed all the expression analyses in R (Script10_Proteomics_Microarrays) with the input files (i) InputFile12_Proteomics1_Microarrays for which a low value was imputed to undetected proteins*; and (ii) InputFile13_Proteomics2_Microarrays corresponding to the raw data, without imputation.
**such imputation is necessary for the statistical analysis of gene expression differences between males and females (Figure 4-source data 1 and 2).*

Obtaining Ka & Ks values for *Schistosoma*

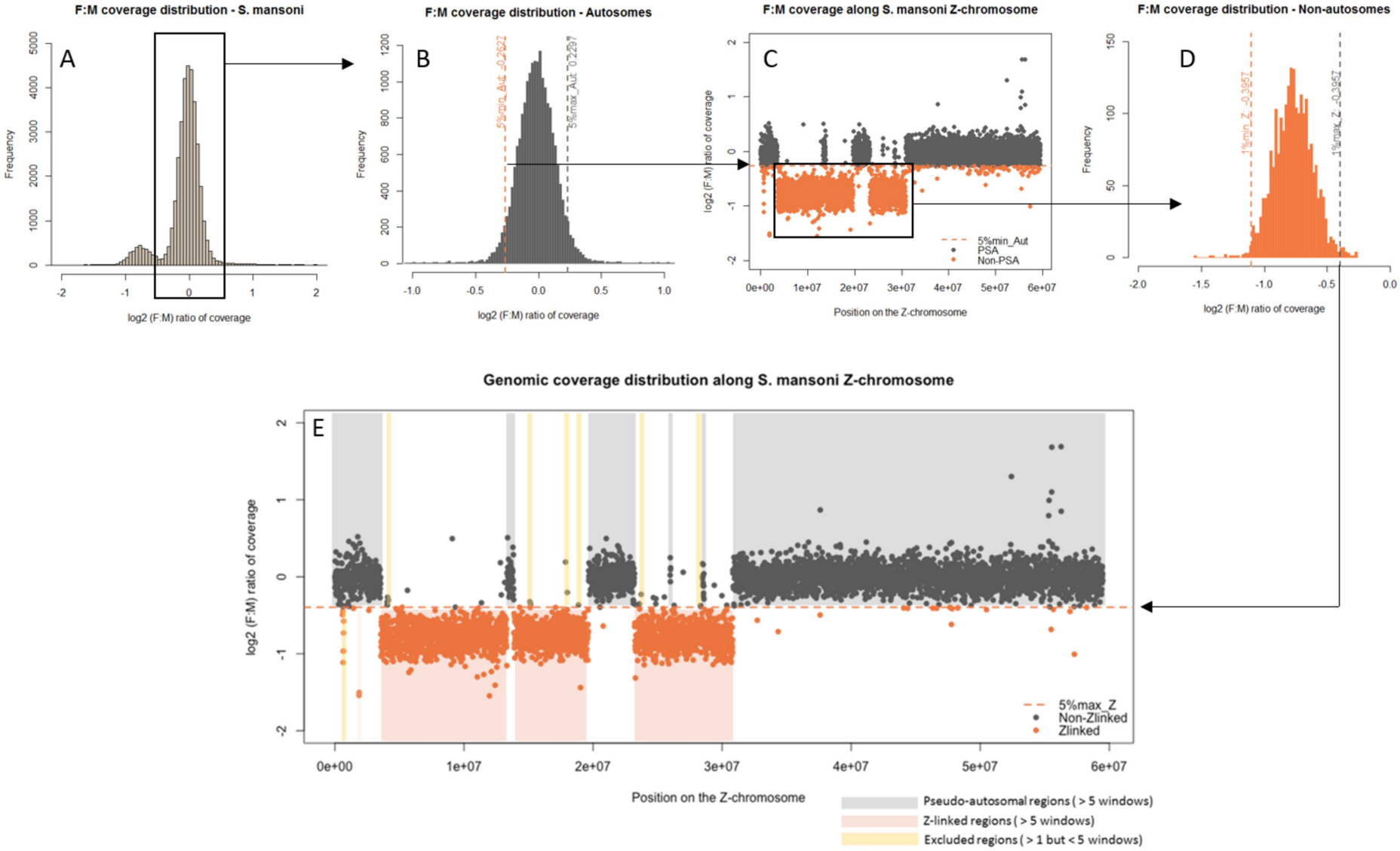
CDS sequences were downloaded from Parasite Wormbase, and 1:1 orthologs detected with a stringent reciprocal best hit approach (Blat with a translated query and database, and a minimum mapping score of 100). These orthologs were aligned with TranslatorX (http://translatorx.co.uk/) with the Gblocks option, and Ka and Ks values were obtained with KaKs_calculator (https://sourceforge.net/projects/kakscalculator2/, model NG). The specific commands used are provided below. Tables of Ka and Ks values are provided as input files 15 and 16.

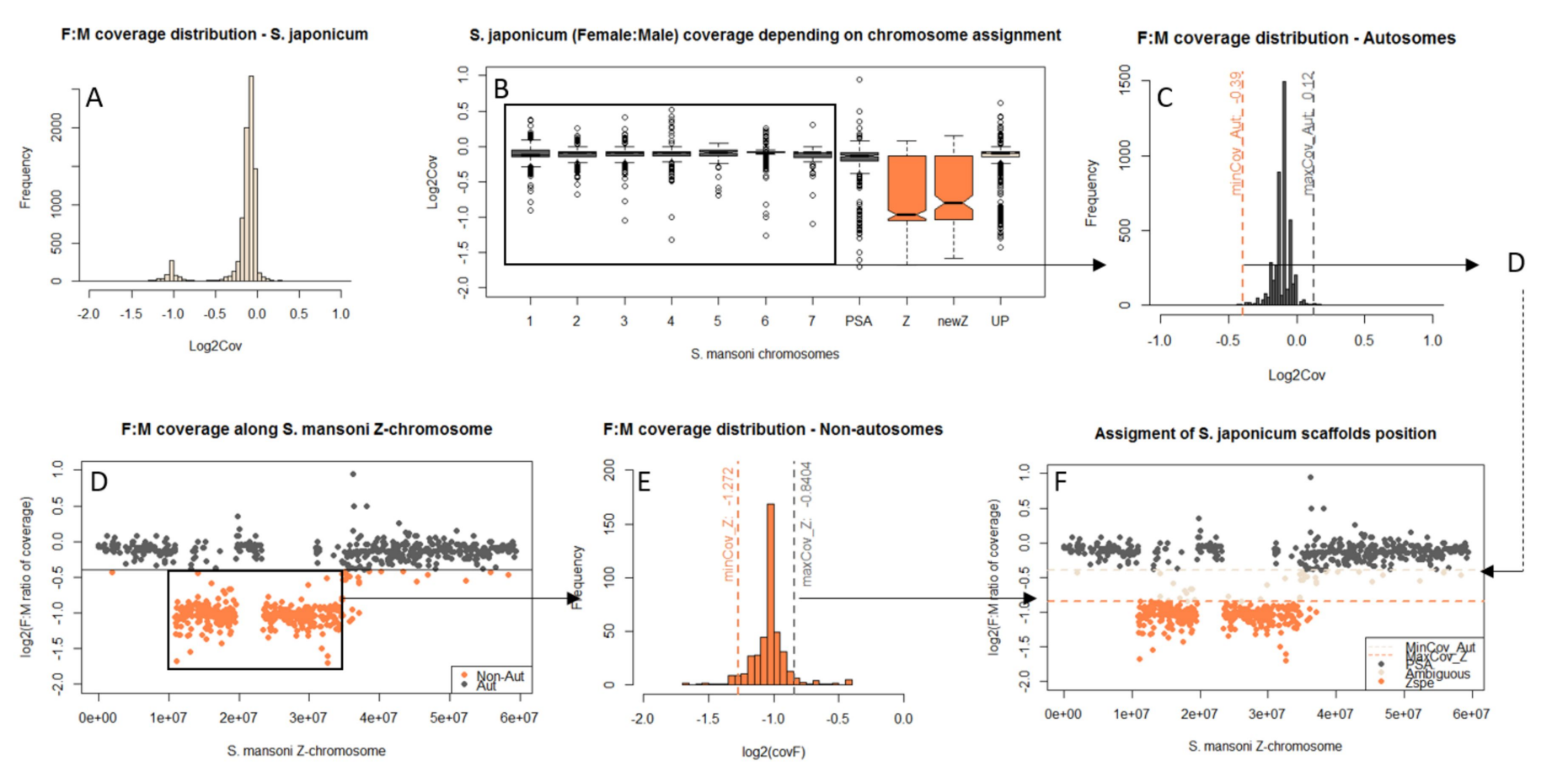
S. mansoni / S. haematobium

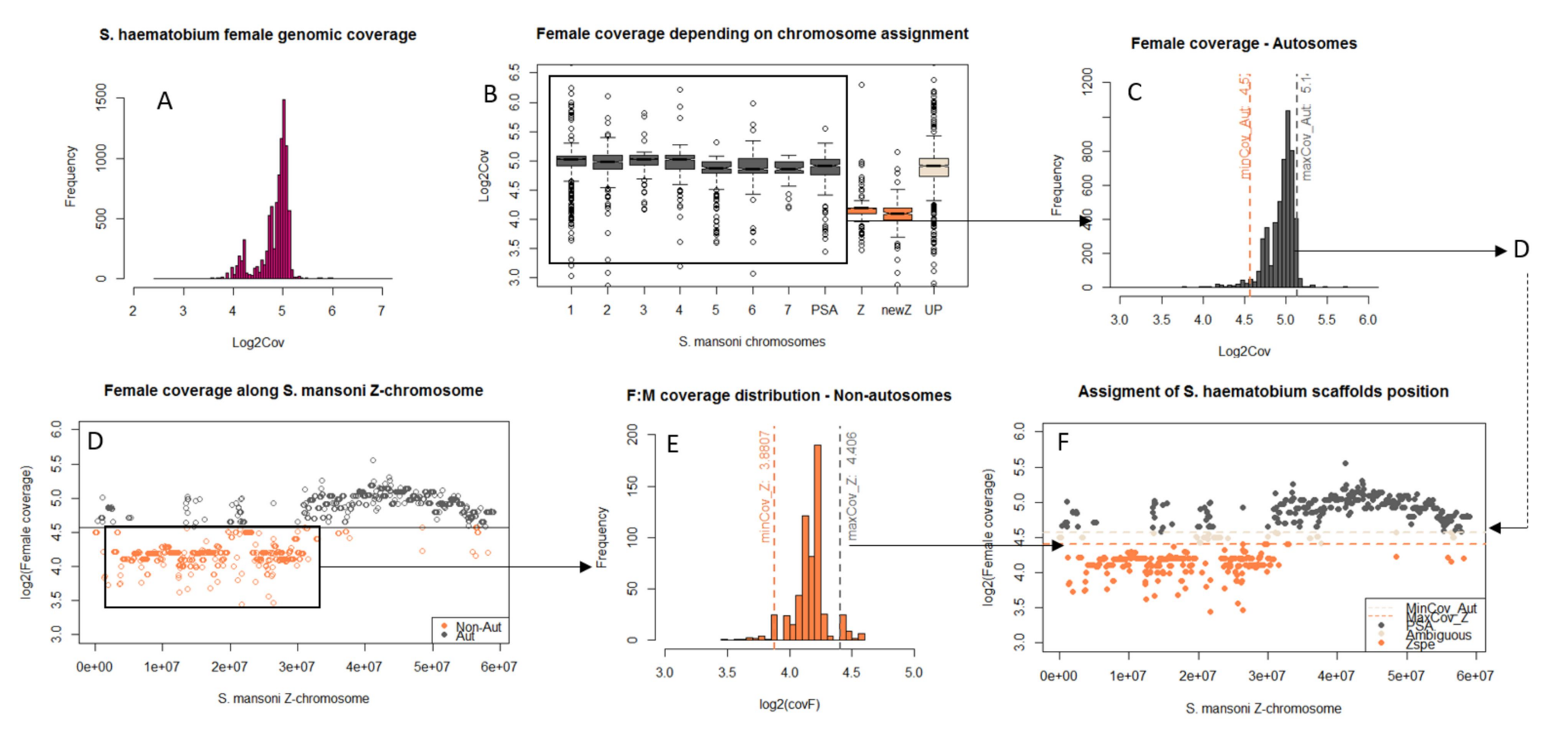
```
# Get CDS

wget
ftp://ftp.ebi.ac.uk/pub/databases/wormbase/parasite/releases/WBPS10/species/schistosoma
_mansoni/PRJEA36577/schistosoma_mansoni.PRJEA36577.WBPS10.CDS_transcripts.fa.gz

gzip -d schistosoma_mansoni.PRJEA36577.WBPS10.CDS_transcripts.fa.gz

mv schistosoma_mansoni.PRJEA36577.WBPS10.CDS_transcripts.fa sman_cds.fa

wget
ftp://ftp.ebi.ac.uk/pub/databases/wormbase/parasite/releases/WBPS10/species/schistosoma
_haematobium/PRJNA78265/schistosoma_haematobium.PRJNA78265.WBPS10.CDS_transcripts.fa.gz

gzip -d schistosoma_haematobium.PRJNA78265.WBPS10.CDS_transcripts.fa.gz

mv schistosoma_haematobium.PRJNA78265.WBPS10.CDS_transcripts.fa shaem_cds.fa

# Run KaKS pipeline

perl Script12_KaKs.pl Sman_Shaem/shaem_cds.fa Sman_Shaem/sman_cds.fa Sman_Shaem/
```

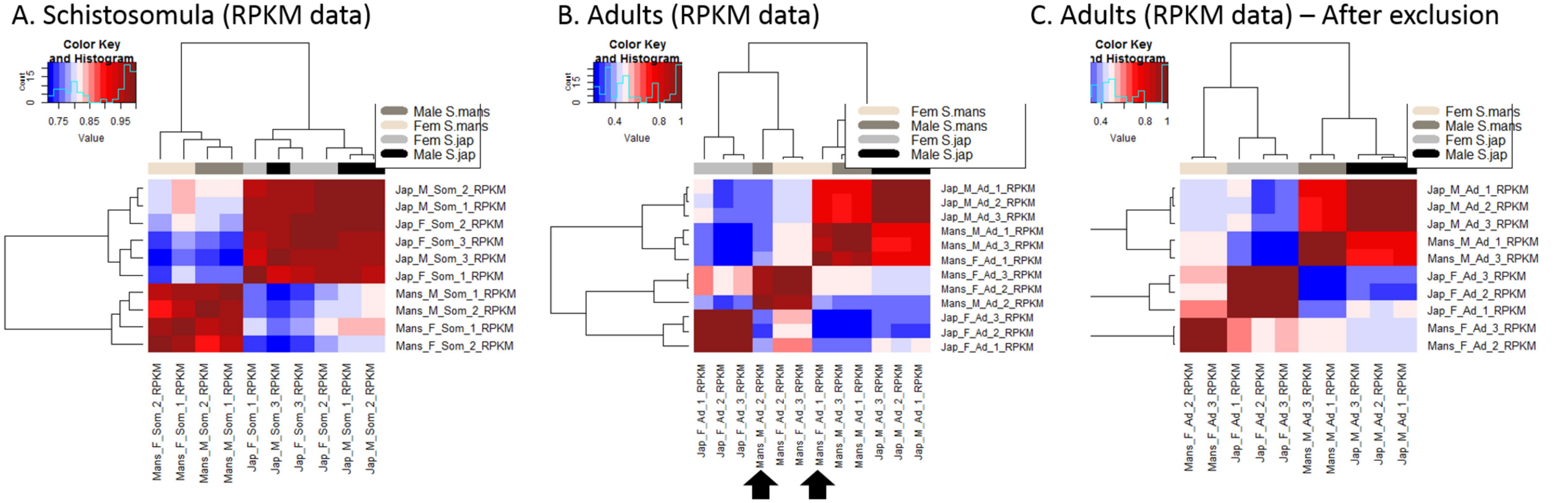
S. mansoni / S. japonicum

```
# Get CDS

wget
ftp://ftp.ebi.ac.uk/pub/databases/wormbase/parasite/releases/WBPS10/species/schistosoma
_mansoni/PRJEA36577/schistosoma_mansoni.PRJEA36577.WBPS10.CDS_transcripts.fa.gz

gzip -d schistosoma_mansoni.PRJEA36577.WBPS10.CDS_transcripts.fa.gz

wget
ftp://ftp.ebi.ac.uk/pub/databases/wormbase/parasite/releases/WBPS10/species/schistosoma
_japonicum/PRJEA34885/schistosoma_japonicum.PRJEA34885.WBPS10.CDS_transcripts.fa.gz

gzip -d schistosoma_japonicum.PRJEA34885.WBPS10.CDS_transcripts.fa.gz


mv schistosoma_japonicum.PRJEA34885.WBPS10.CDS_transcripts.fa sjap_cds.fa
mv schistosoma_mansoni.PRJEA36577.WBPS10.CDS_transcripts.fa sman_cds.fa

# Run KaKS pipeline

perl Script12_KaKs.pl Sman_Sjap/sjap_cds.fa Sman_Sjap/sman_cds.fa Sman_Sjap/
```

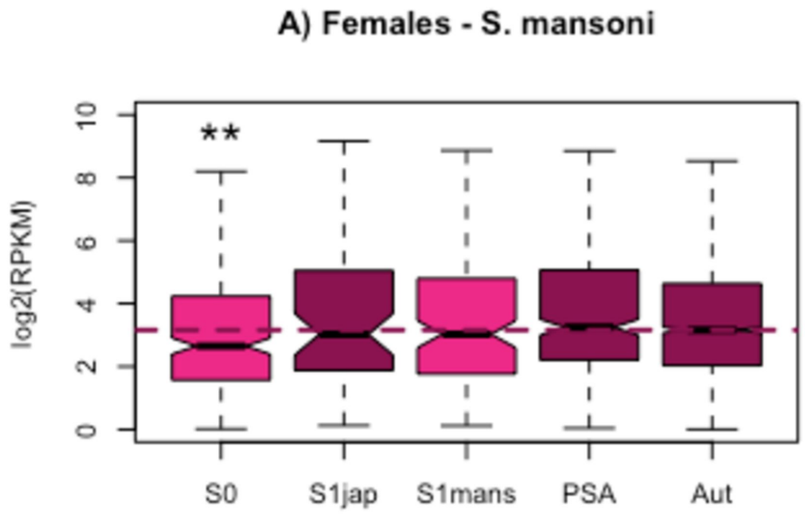

F:M coverage distribution - S. mansoni
A
Frequency
log2 (F:M) ratio of coverage
F:M coverage distribution - Autosomes
B
5%min_Aut -0.2622
5%max_Aut 0.2297
Frequency
log2 (F:M) ratio of coverage
F:M coverage along S. mansoni Z-chromosome
C
log2 (F:M) ratio of coverage
5%min_Aut
PSA
Non-PSA
Position on the Z-chromosome
F:M coverage distribution - Non-autosomes
D
1%min_Z -0.3957
1%max_Z -0.3957
Frequency
log2 (F:M) ratio of coverage
Genomic coverage distribution along S. mansoni Z-chromosome
E
log2 (F:M) ratio of coverage
5%max_Z
Non-Zlinked
Zlinked
Position on the Z-chromosome
Pseudo-autosomal regions (> 5 windows)
Z-linked regions (> 5 windows)
Excluded regions (> 1 but < 5 windows)

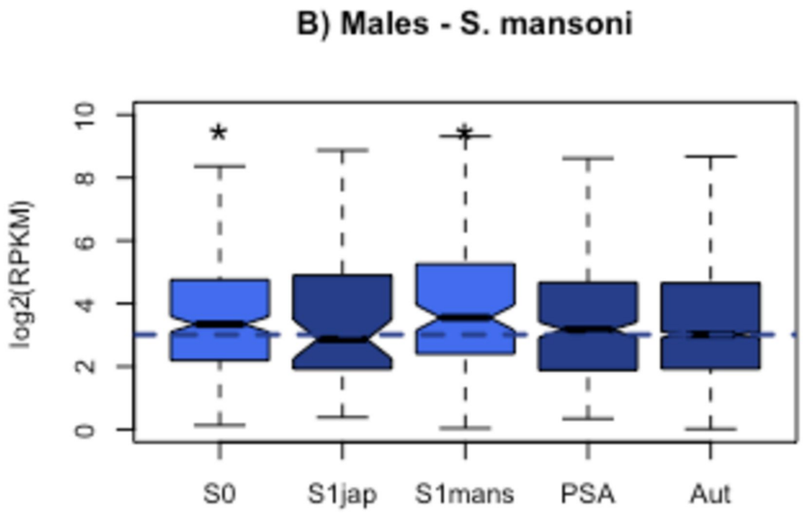

F:M coverage distribution - S. japonicum
A
Frequency
Log2Cov
S. japonicum (Female:Male) coverage depending on chromosome assignment
B
Log2Cov
1
2
3
4
5
6
7
PSA
Z
newZ
UP
S. mansoni chromosomes
F:M coverage distribution - Autosomes
C
minCov_Aut: -0.39
maxCov_Aut: 0.12
Frequency
Log2Cov
D
F:M coverage along S. mansoni Z-chromosome
D
log2(F:M ratio of coverage)
Non-Aut
Aut
S. mansoni Z-chromosome
F:M coverage distribution - Non-autosomes
E
minCov_Z: -1.272
maxCov_Z: -0.8404
Frequency
log2(covF)
Assigment of S. japonicum scaffolds position
F
log2(F:M ratio of coverage)
MinCov_Aut
MaxCov_Z
PSA
Ambiguous
Zspe
S. mansoni Z-chromosome

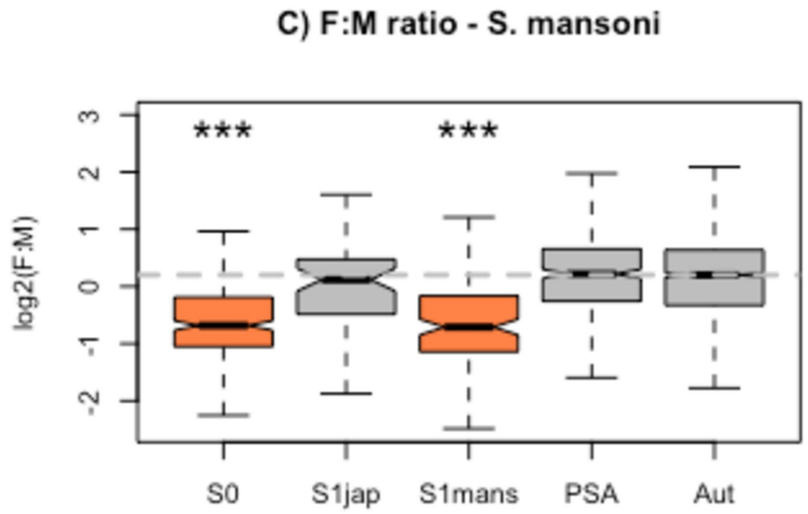

S. haematobium female genomic coverage
A
Frequency
Log2Cov
Female coverage depending on chromosome assignment
B
Log2Cov
1 2 3 4 5 6 7 PSA Z newZ UP
S. mansoni chromosomes
Female coverage - Autosomes
C
minCov_Aut: 4.57
maxCov_Aut: 5.1
Frequency
Log2Cov
D
Female coverage along S. mansoni Z-chromosome
D
log2(Female coverage)
Non-Aut
Aut
S. mansoni Z-chromosome
F:M coverage distribution - Non-autosomes
E
minCov_Z: 3.8807
maxCov_Z: 4.406
Frequency
log2(covF)
Assigment of S. haematobium scaffolds position
F
log2(Female coverage)
MinCov_Aut
MaxCov_Z
PSA
Ambiguous
Zspe
S. mansoni Z-chromosome

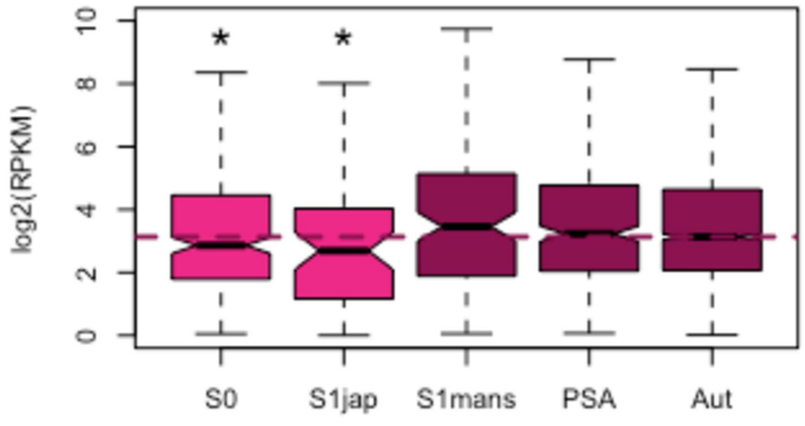

A. Schistosomula (RPKM data)
Color Key and Histogram
Count
0
15
0.75
0.85
0.95
Value
Male S.mans
Fem S.mans
Fem S.jap
Male S.jap
Jap_M_Som_2_RPKM
Jap_M_Som_1_RPKM
Jap_F_Som_2_RPKM
Jap_F_Som_3_RPKM
Jap_M_Som_3_RPKM
Jap_F_Som_1_RPKM
Mans_M_Som_1_RPKM
Mans_M_Som_2_RPKM
Mans_F_Som_1_RPKM
Mans_F_Som_2_RPKM
Mans_F_Som_2_RPKM
Mans_F_Som_1_RPKM
Mans_M_Som_2_RPKM
Mans_M_Som_1_RPKM
Jap_F_Som_1_RPKM
Jap_M_Som_3_RPKM
Jap_F_Som_3_RPKM
Jap_F_Som_2_RPKM
Jap_M_Som_1_RPKM
Jap_M_Som_2_RPKM
B. Adults (RPKM data)
Color Key and Histogram
Count
0
15
0.4
0.8
1
Value
Fem S.mans
Male S.mans
Fem S.jap
Male S.jap
Jap_M_Ad_1_RPKM
Jap_M_Ad_2_RPKM
Jap_M_Ad_3_RPKM
Mans_M_Ad_1_RPKM
Mans_M_Ad_3_RPKM
Mans_F_Ad_1_RPKM
Mans_F_Ad_3_RPKM
Mans_F_Ad_2_RPKM
Mans_M_Ad_2_RPKM
Jap_F_Ad_3_RPKM
Jap_F_Ad_2_RPKM
Jap_F_Ad_1_RPKM
Jap_F_Ad_1_RPKM
Jap_F_Ad_2_RPKM
Jap_F_Ad_3_RPKM
Mans_M_Ad_2_RPKM
Mans_M_Ad_1_RPKM
Mans_M_Ad_3_RPKM
Mans_F_Ad_2_RPKM
Mans_F_Ad_3_RPKM
Mans_F_Ad_1_RPKM
Mans_M_Ad_3_RPKM
Mans_M_Ad_1_RPKM
Jap_M_Ad_3_RPKM
Jap_M_Ad_2_RPKM
Jap_M_Ad_1_RPKM
C. Adults (RPKM data) – After exclusion
Color Key and Histogram
0.4
0.8
1
Value
Fem S.mans
Male S.mans
Fem S.jap
Male S.jap
Jap_M_Ad_1_RPKM
Jap_M_Ad_2_RPKM
Jap_M_Ad_3_RPKM
Mans_M_Ad_1_RPKM
Mans_M_Ad_3_RPKM
Jap_F_Ad_3_RPKM
Jap_F_Ad_2_RPKM
Jap_F_Ad_1_RPKM
Mans_F_Ad_3_RPKM
Mans_F_Ad_2_RPKM
Mans_F_Ad_2_RPKM
Mans_F_Ad_3_RPKM
Jap_F_Ad_1_RPKM
Jap_F_Ad_2_RPKM
Jap_F_Ad_3_RPKM
Mans_M_Ad_3_RPKM
Mans_M_Ad_1_RPKM
Jap_M_Ad_3_RPKM
Jap_M_Ad_2_RPKM
Jap_M_Ad_1_RPKM

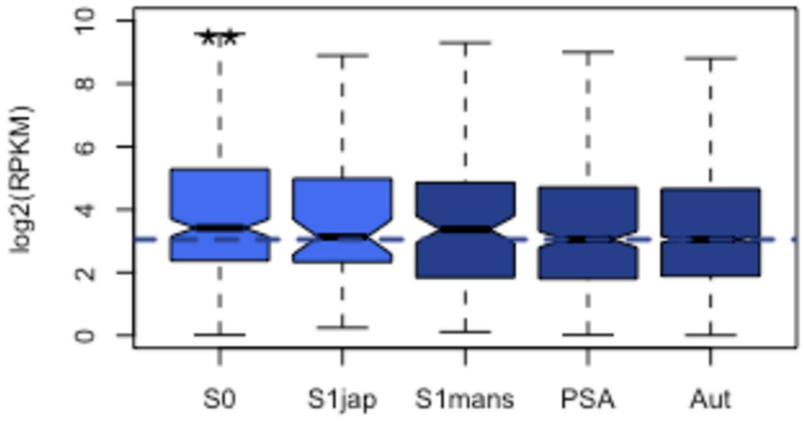
ADULTS - RPKM>1 - Strata Method2 (Exhaustive)

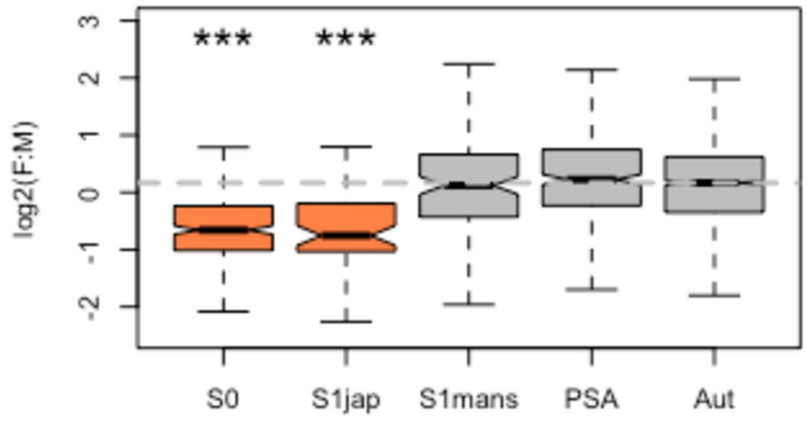


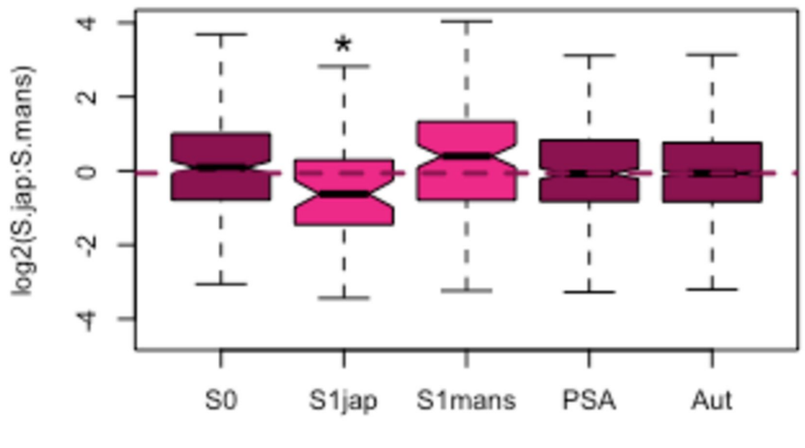


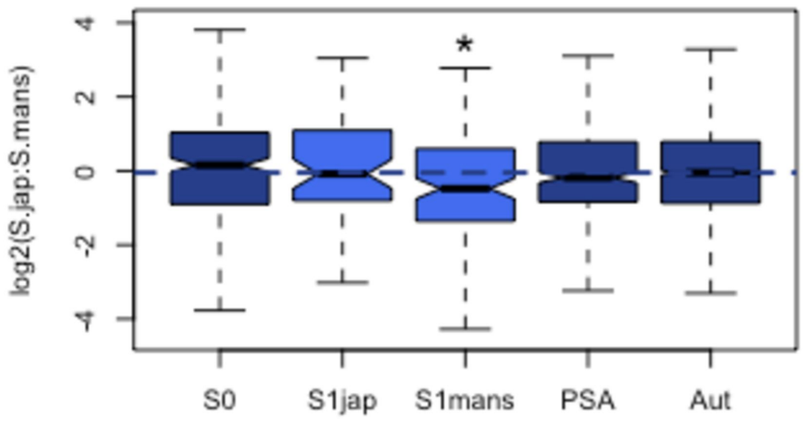


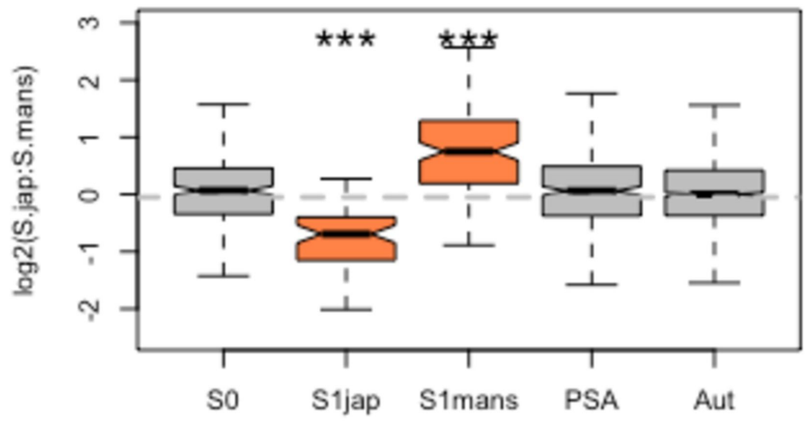




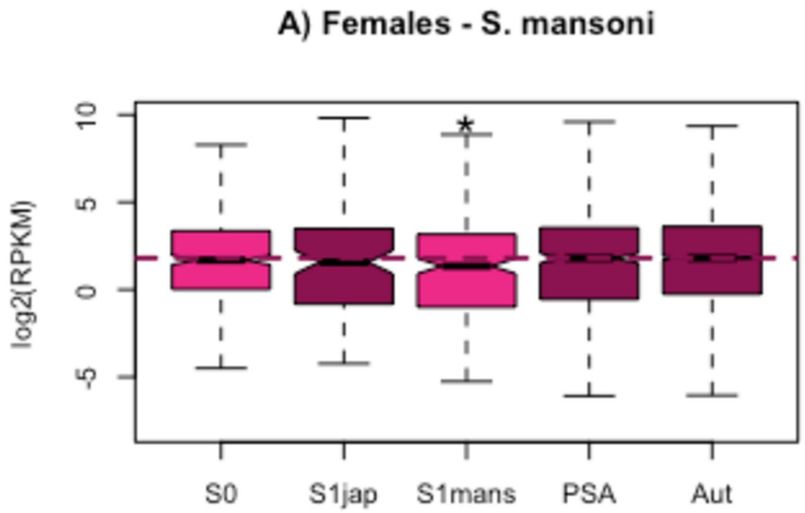


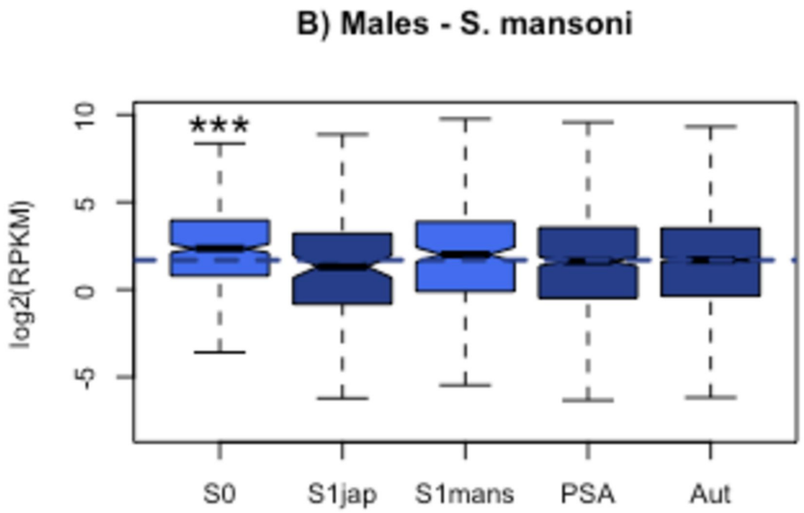


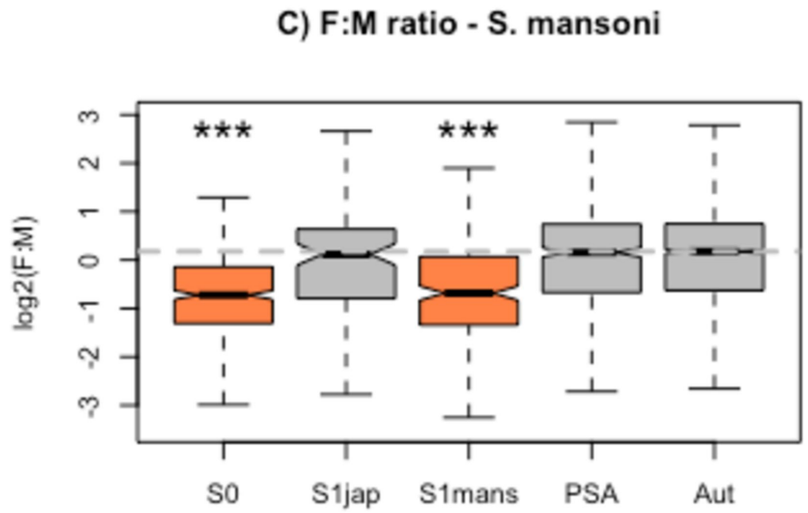


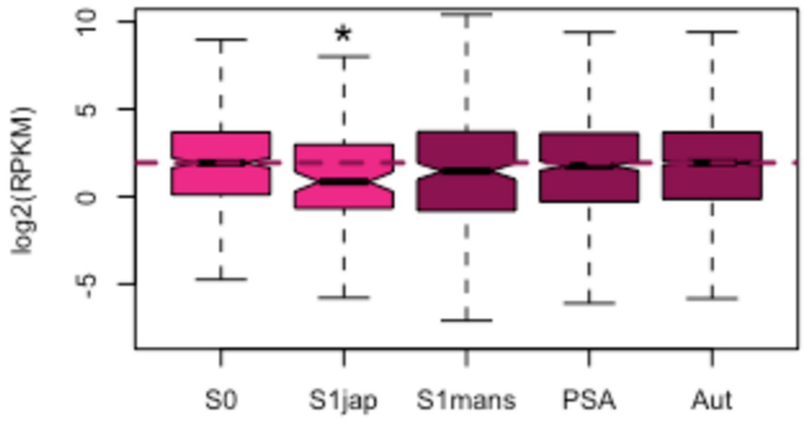


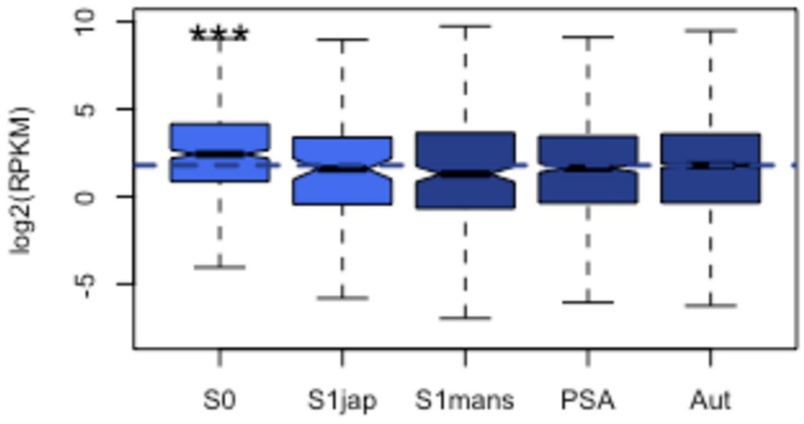
ADULTS - RPKM>0 - Strata Method2 (Exhaustive)

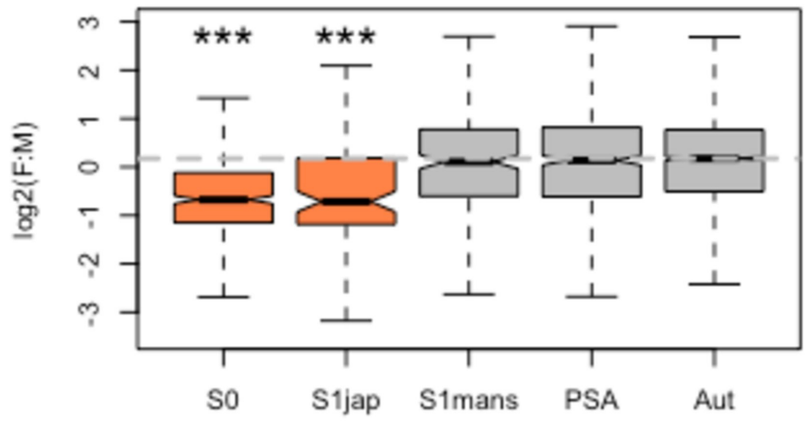


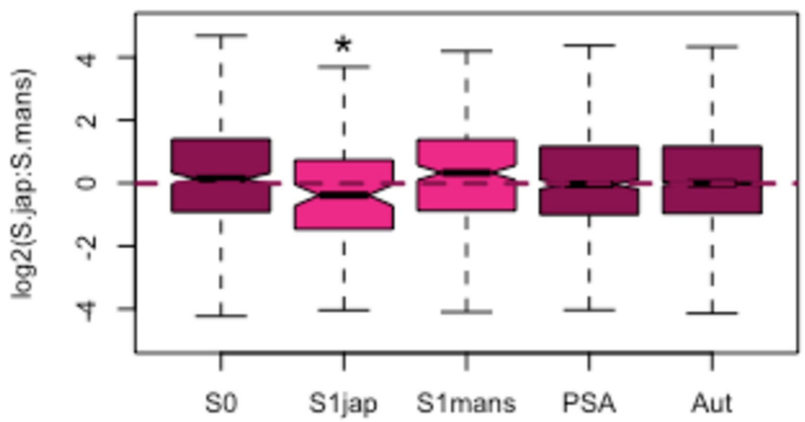


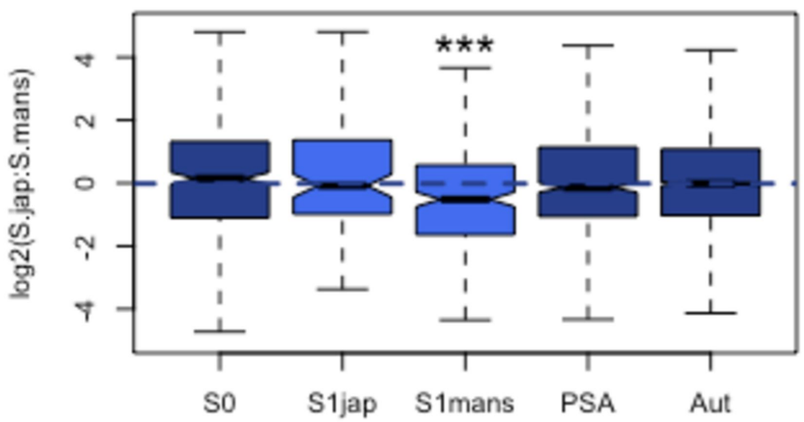


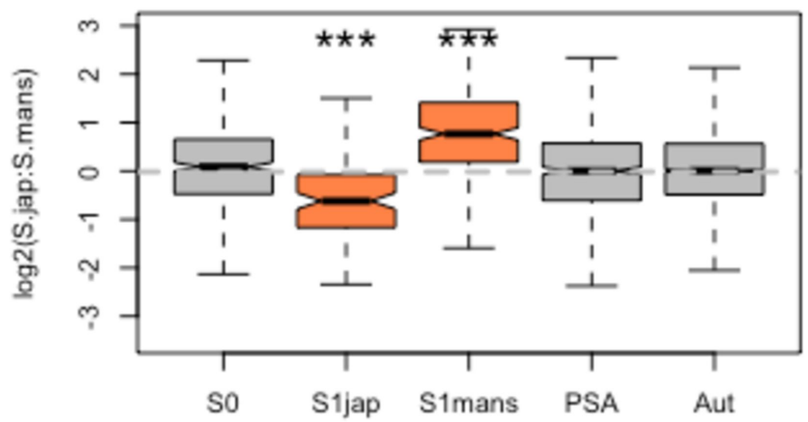






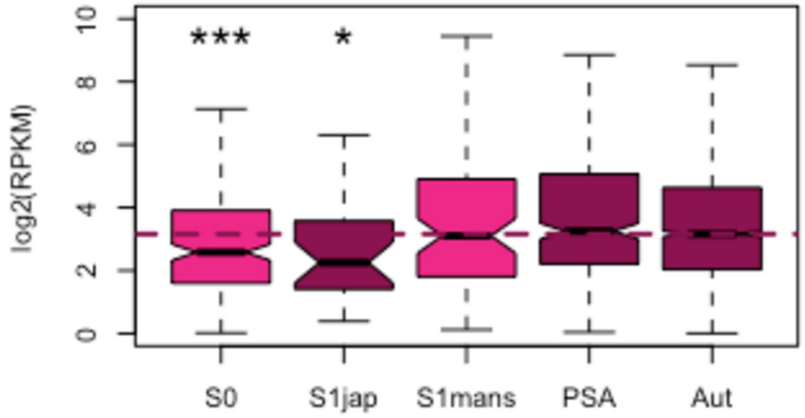




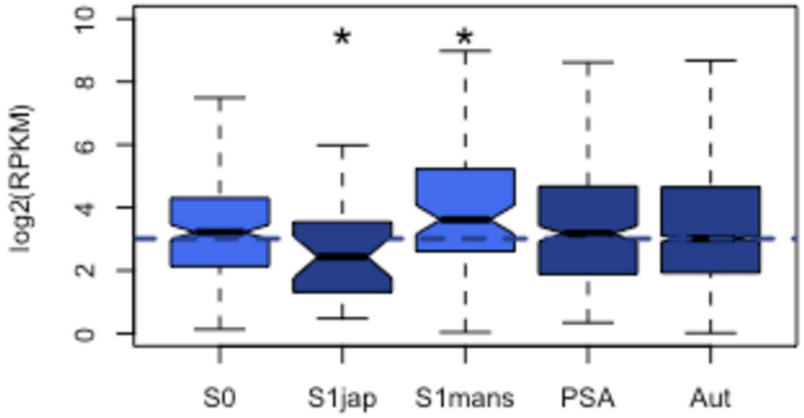


ADULTS - RPKM>1 - Strata Method1 (Stringent)

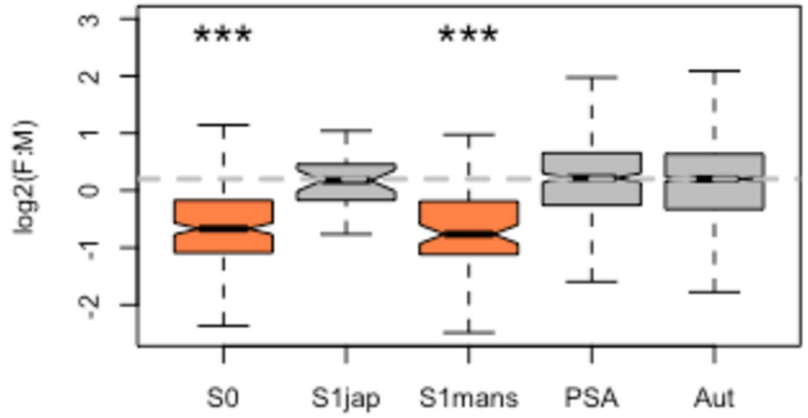
A) Females - S. mansoni



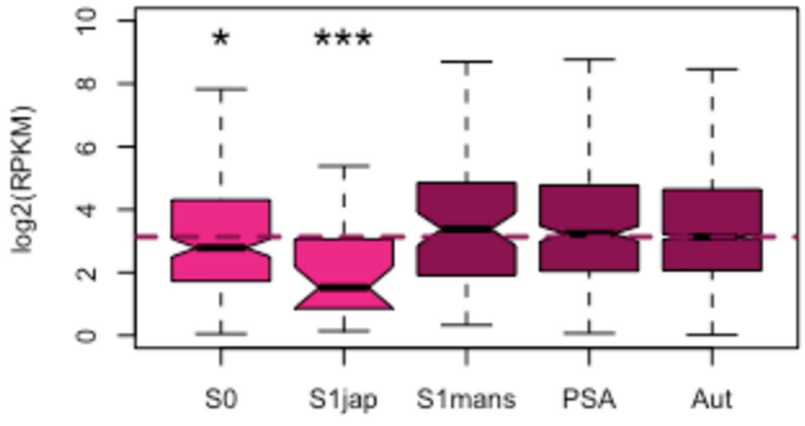
B) Males - S. mansoni



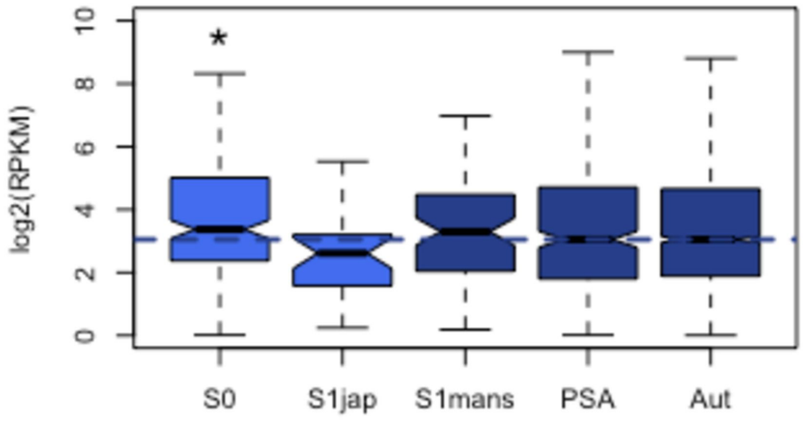
C) F:M ratio - S. mansoni



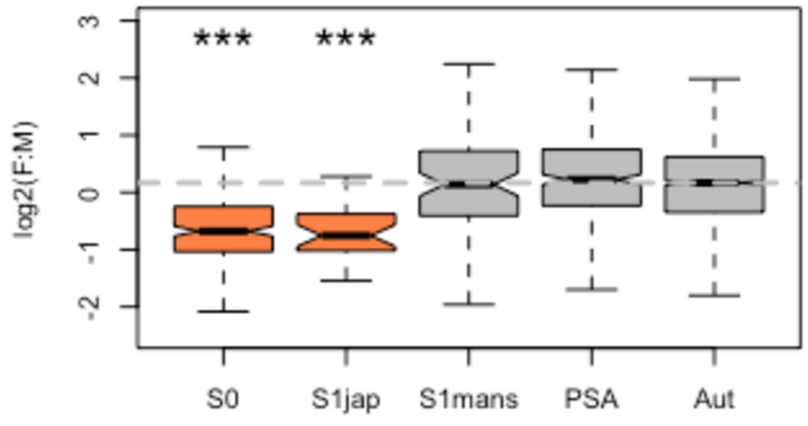
D) Females - S. japonicum



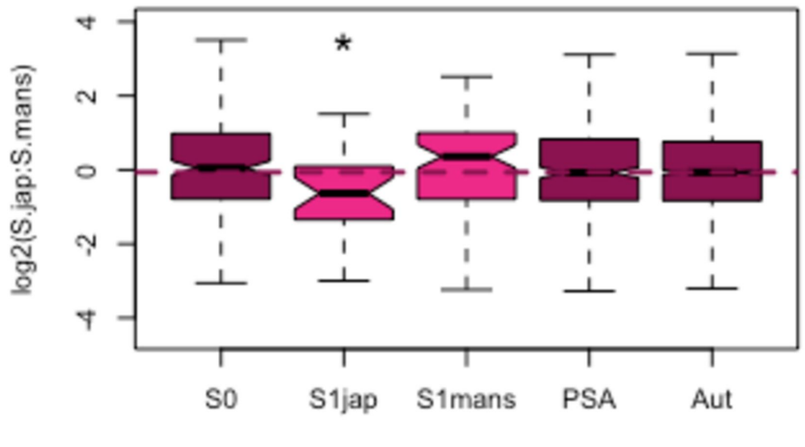
E) Males - S. japonicum



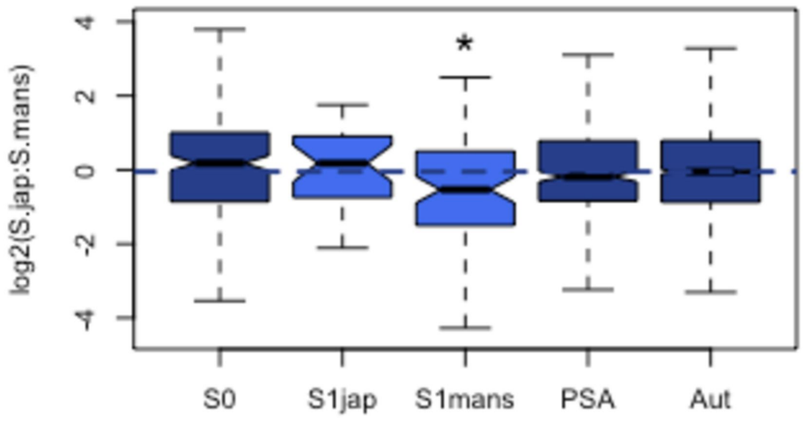
F) F:M ratio - S. japonicum



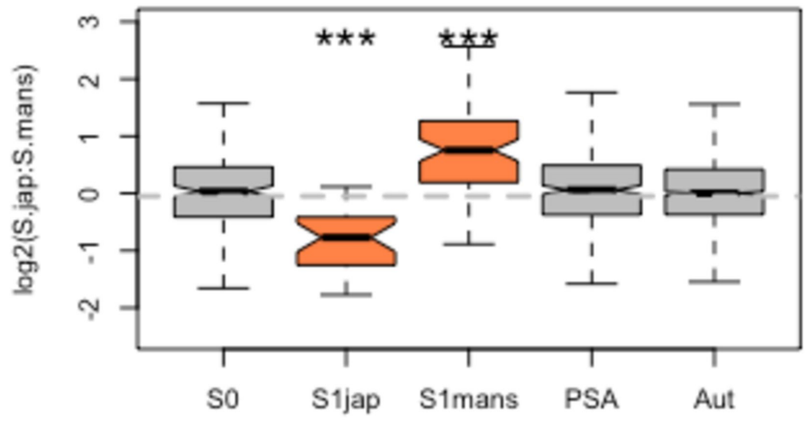
G) Females - S.jap:S.mans



H) Males - S.jap:S.mans

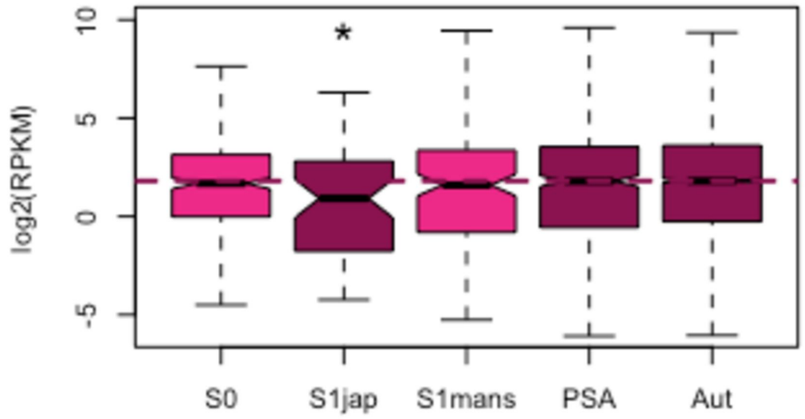


I) F:M ratio - S.jap:S.mans

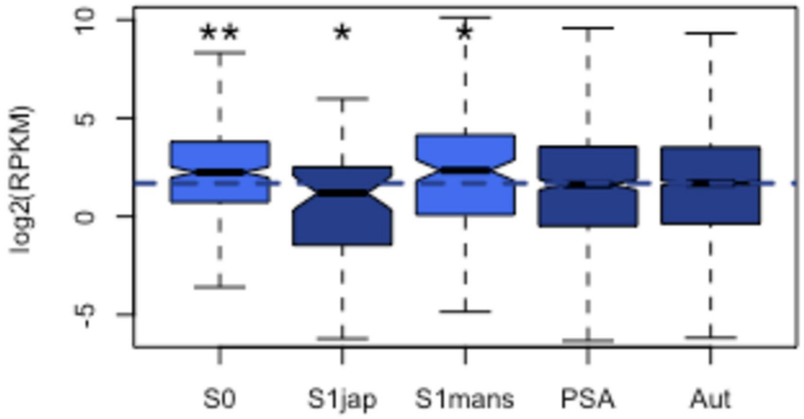


ADULTS - RPKM>0 - Strata Method1 (Stringent)

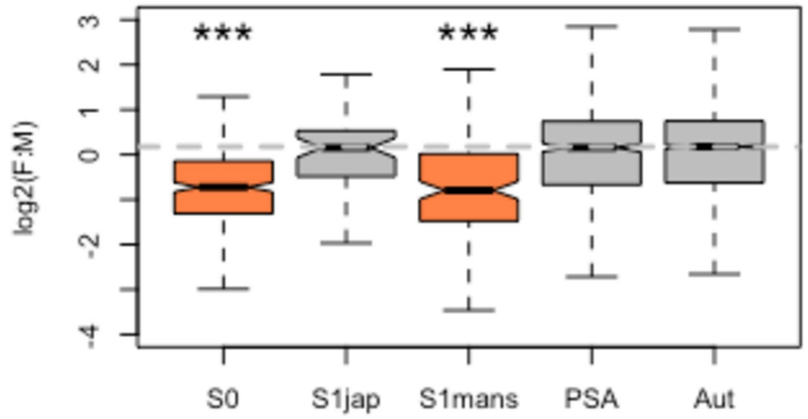
A) Females - S. mansoni



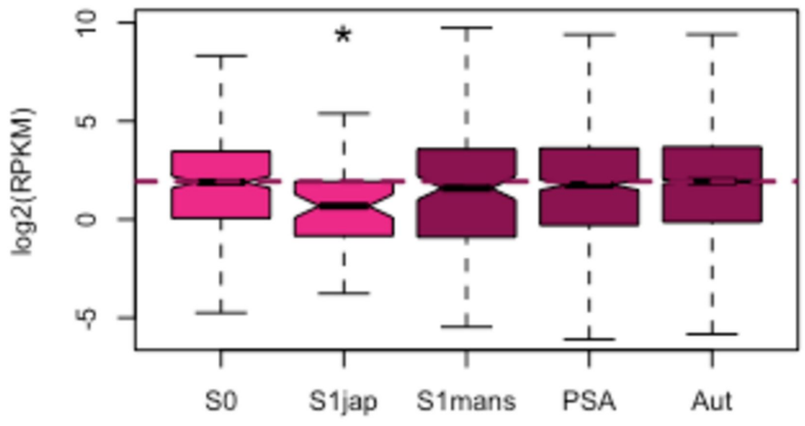
B) Males - S. mansoni



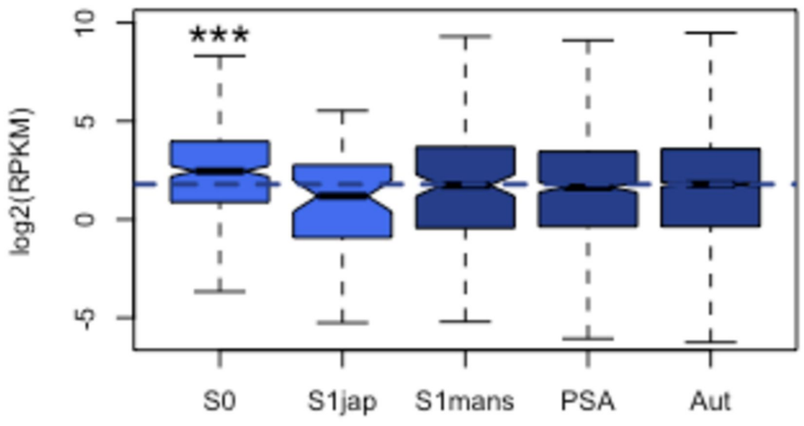
C) F:M ratio - S. mansoni



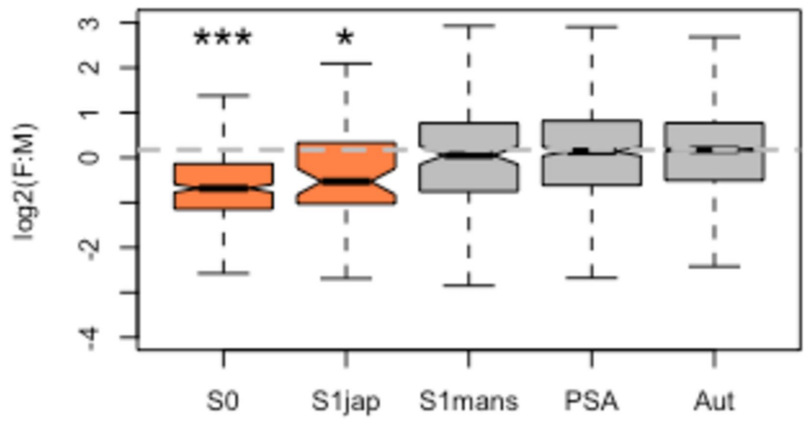
D) Females - S. japonicum



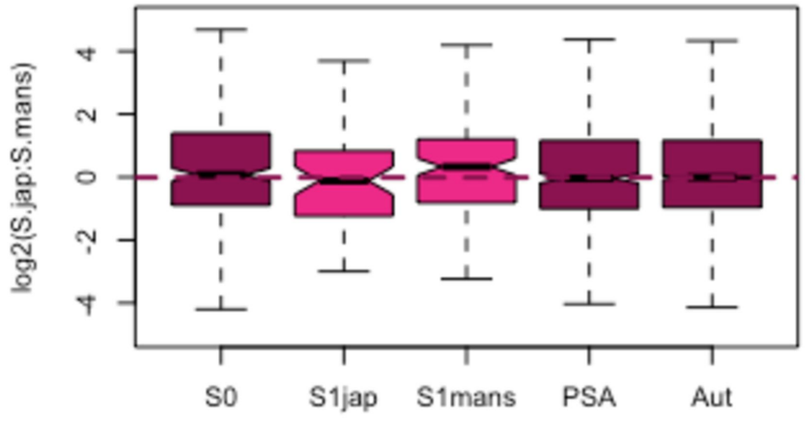
E) Males - S. japonicum



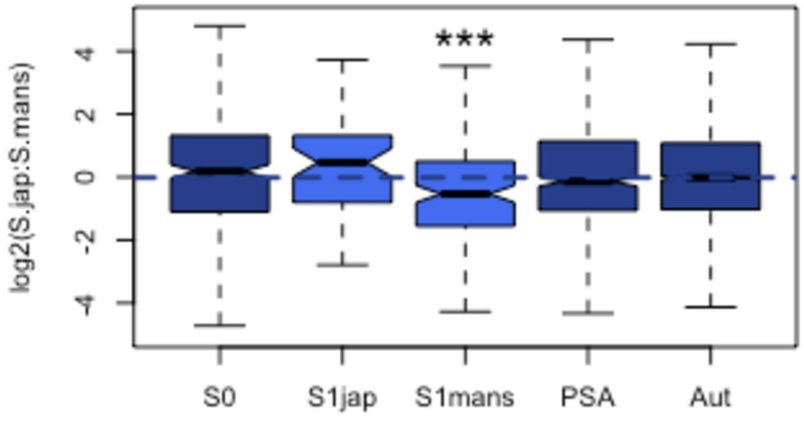
F) F:M ratio - S. japonicum



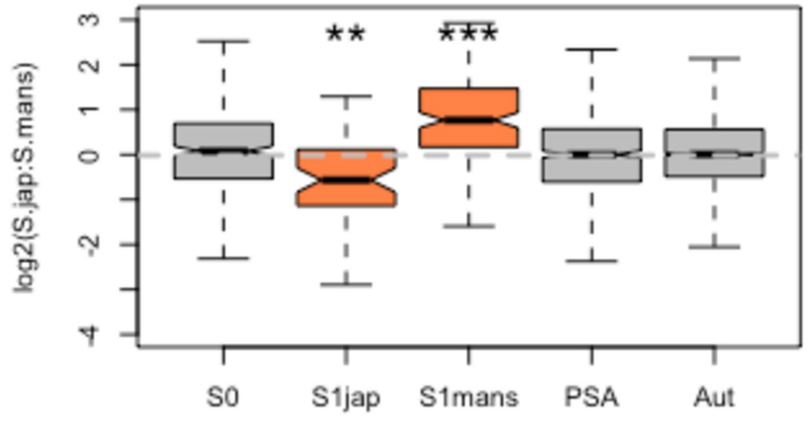
G) Females - S.jap:S.mans



H) Males - S.jap:S.mans

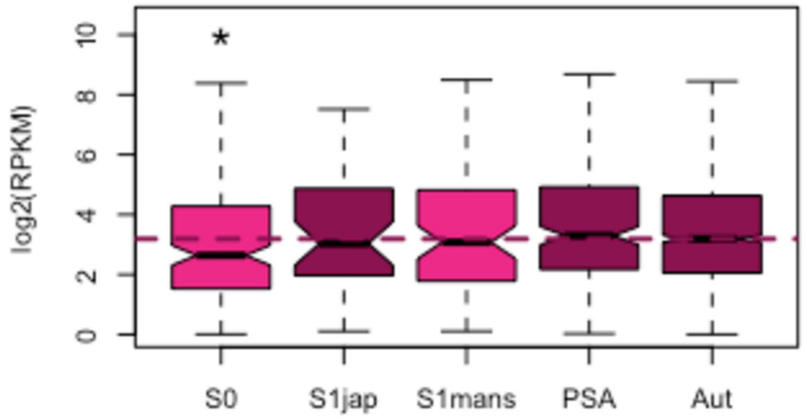


I) F:M ratio - S.jap:S.mans

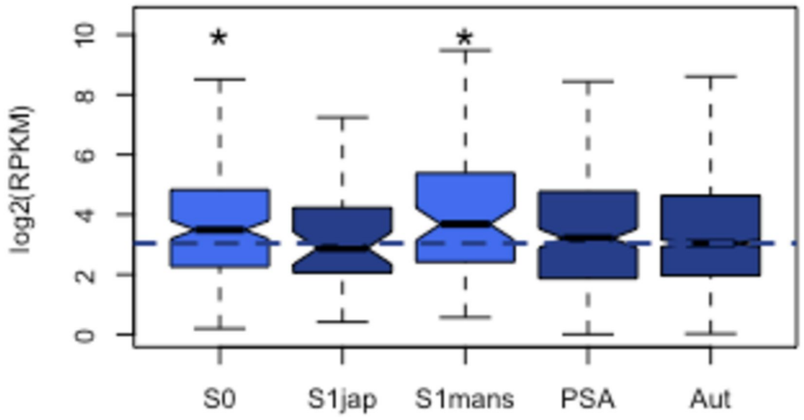


ADULTS - RPKM>1 - Strata Method2 (Exhaustive) - Orthologs Wormbase

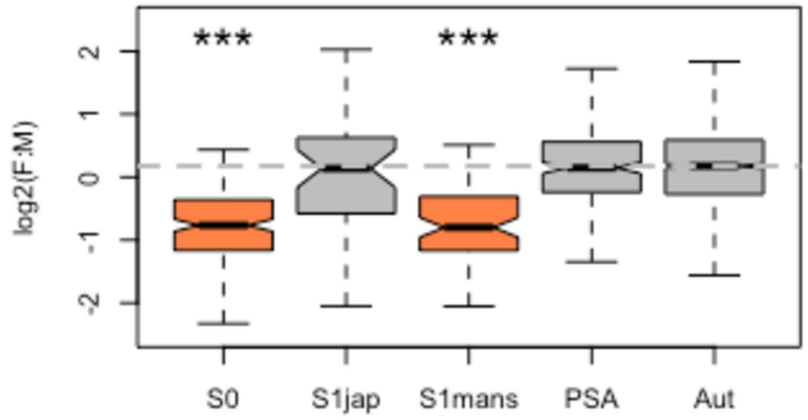
A) Females - S. mansoni



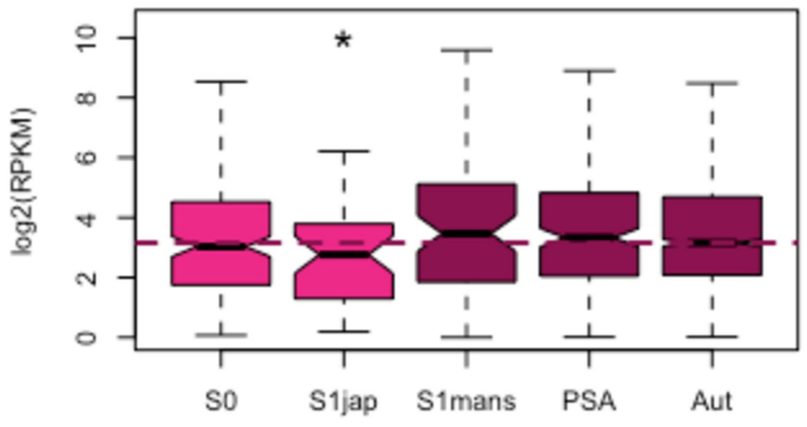
B) Males - S. mansoni



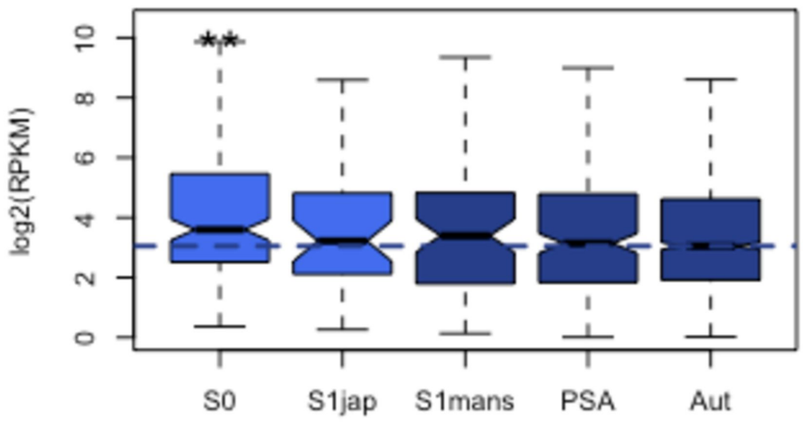
C) F:M ratio - S. mansoni



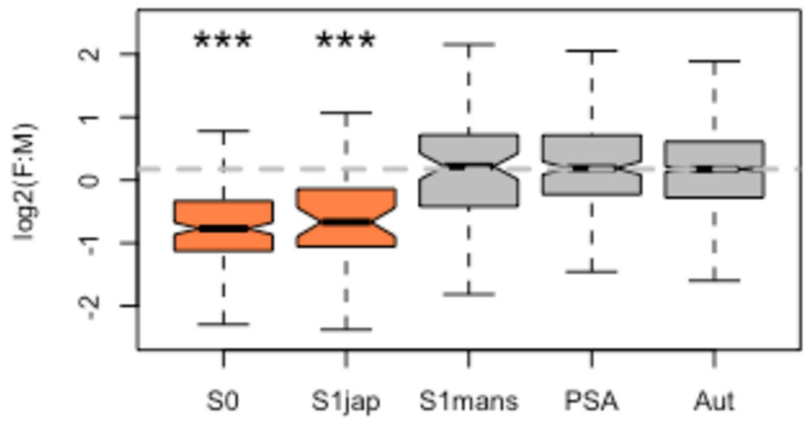
D) Females - S. japonicum



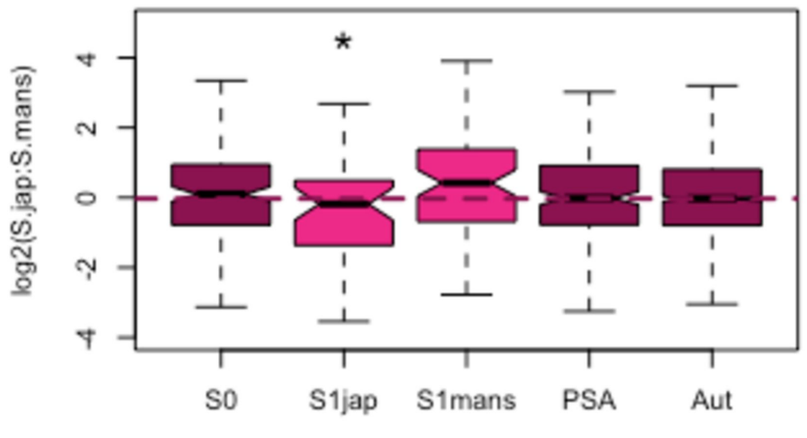
E) Males - S. japonicum



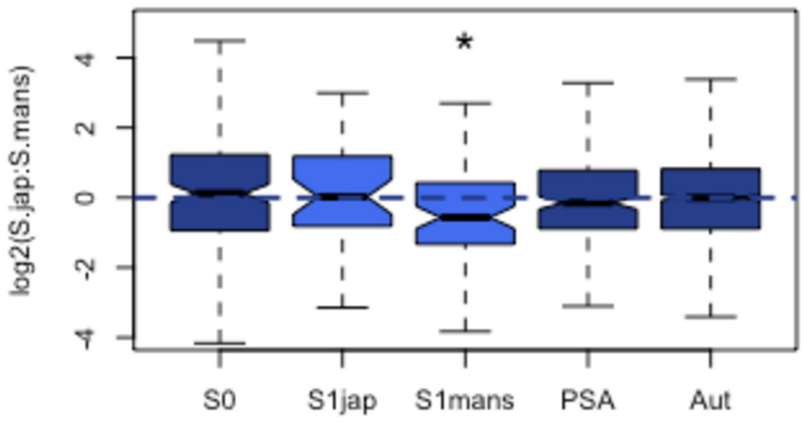
F) F:M ratio - S. japonicum



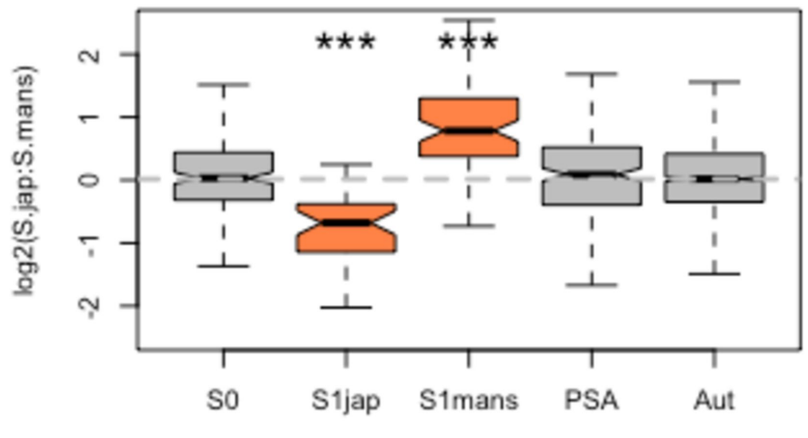
G) Females - S.jap:S.mans



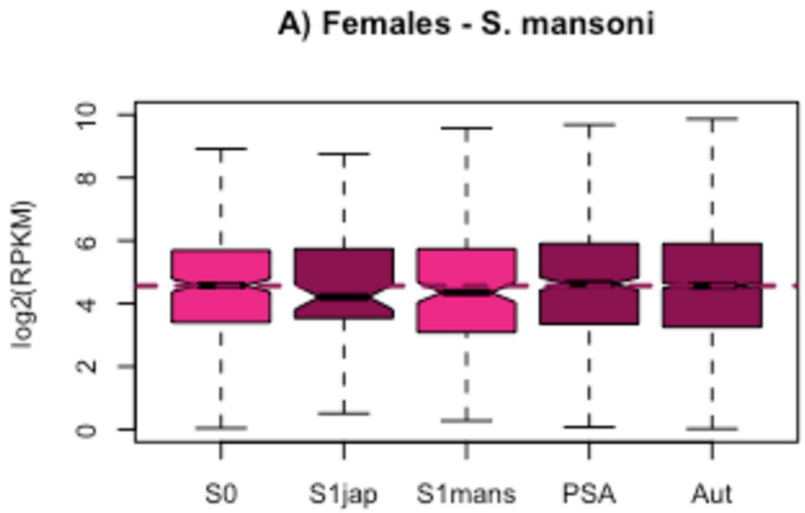
H) Males - S.jap:S.mans

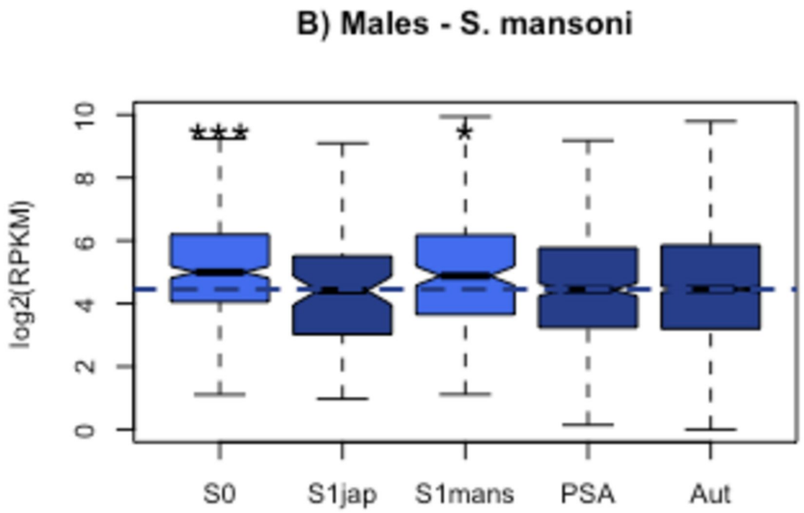


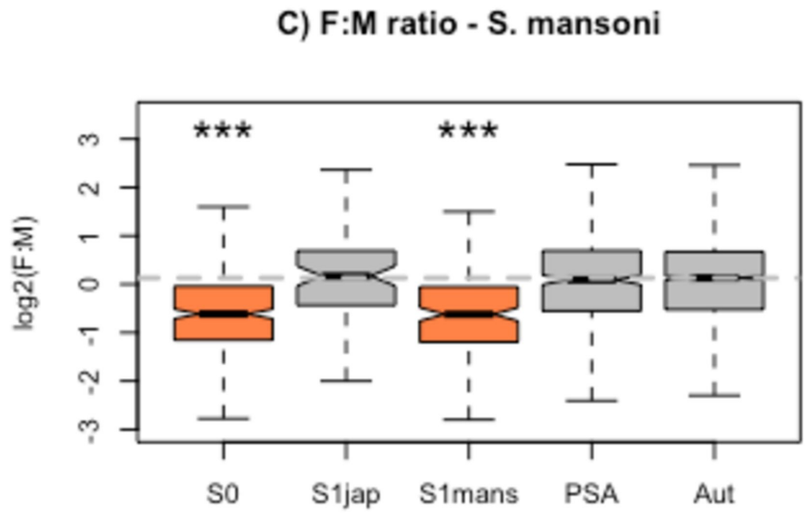
I) F:M ratio - S.jap:S.mans

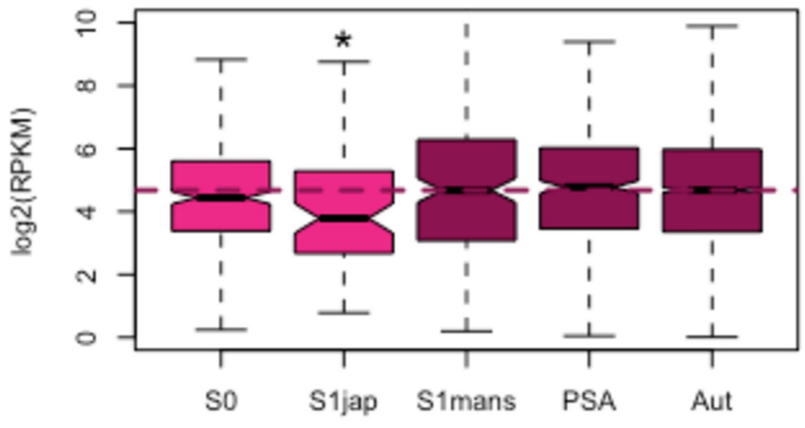


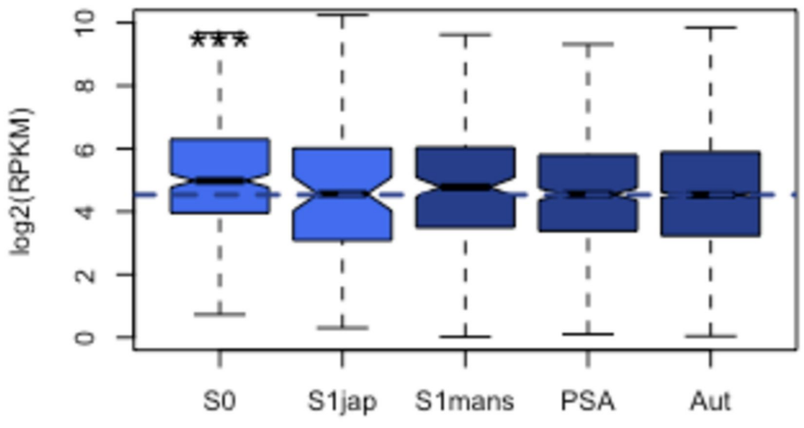
ADULTS - TPM>1 - Strata Method2 (Exhaustive)

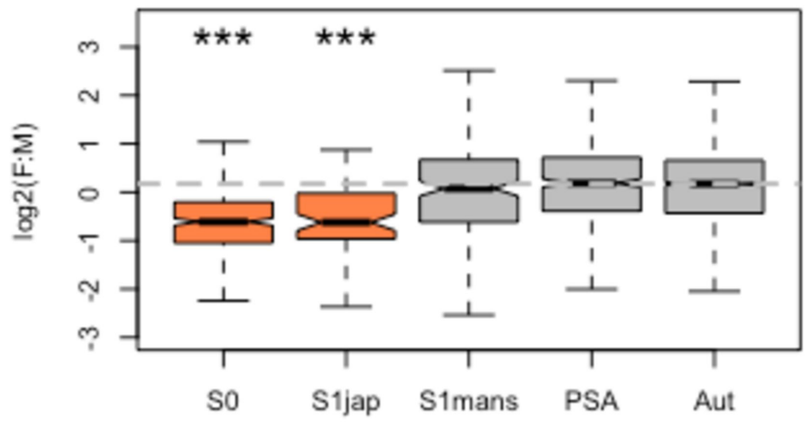


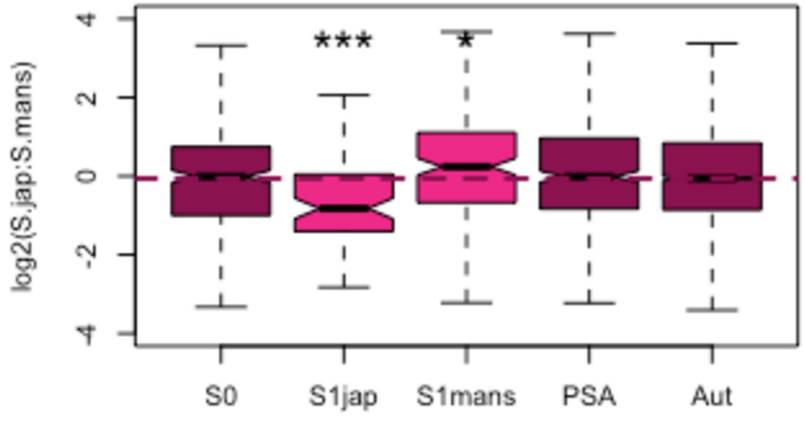


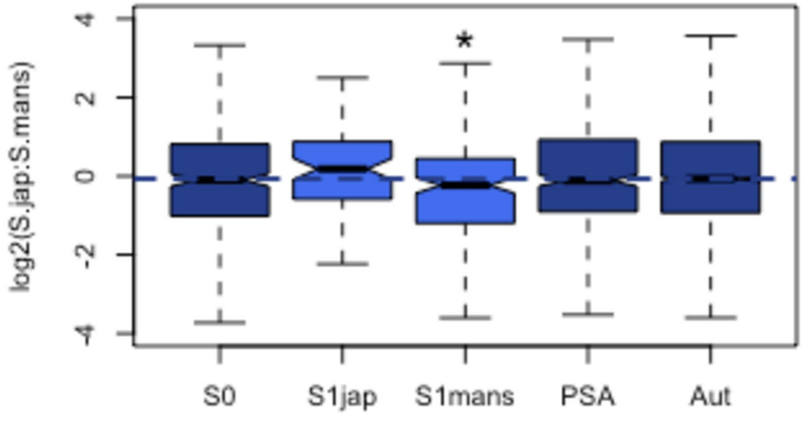


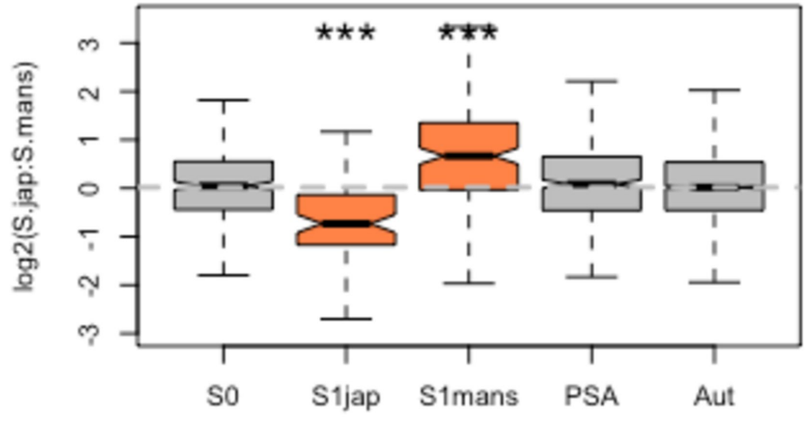




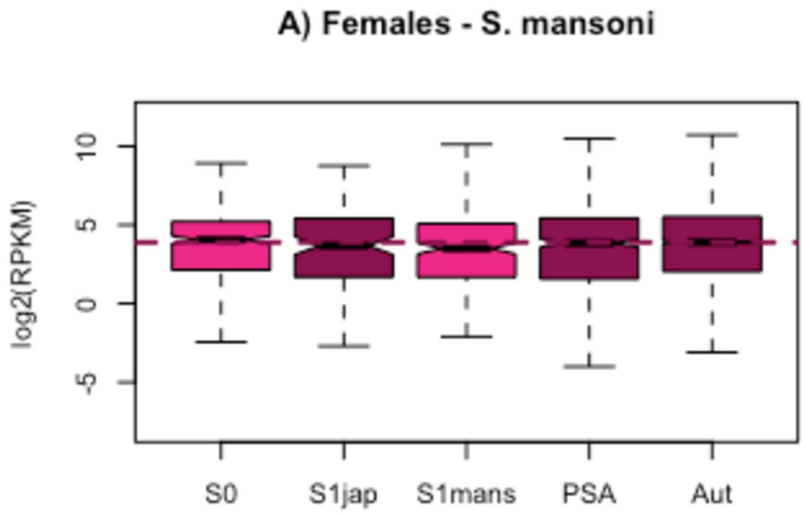


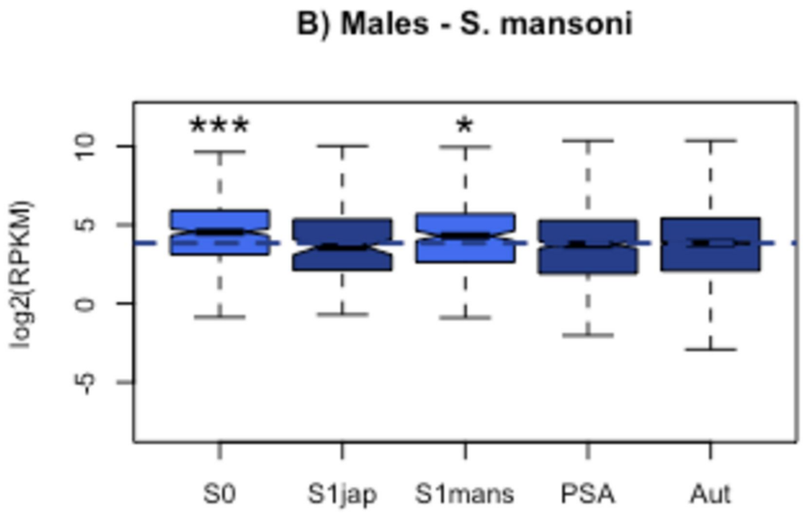


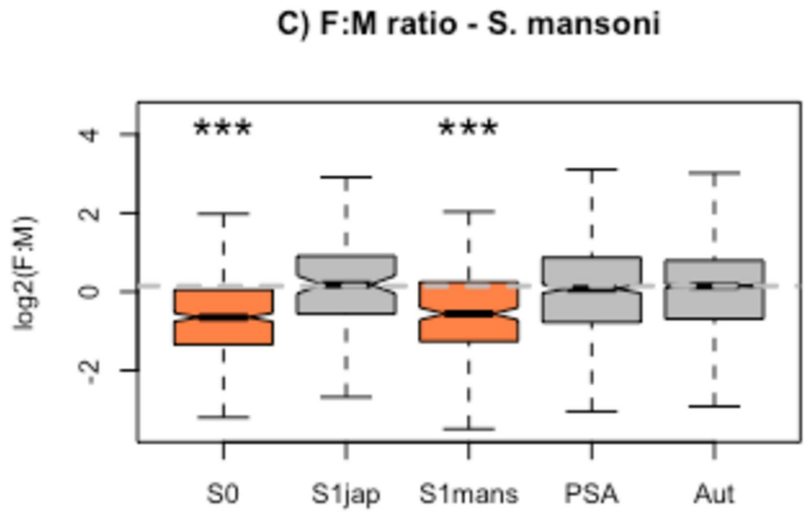


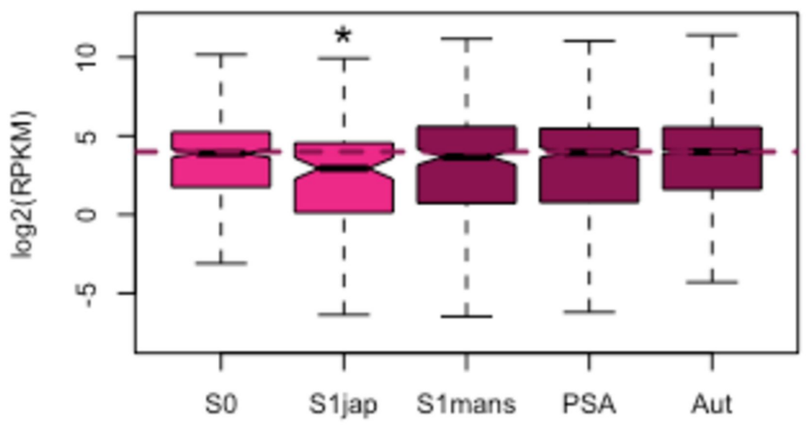


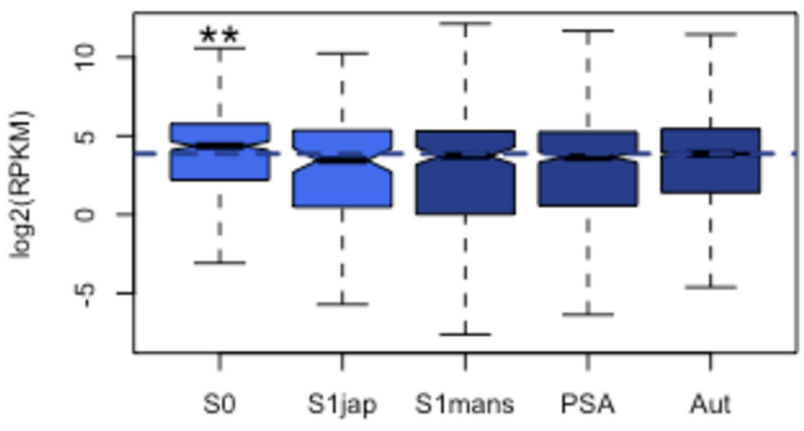
ADULTS - TPM>0 - Strata Method2 (Exhaustive)

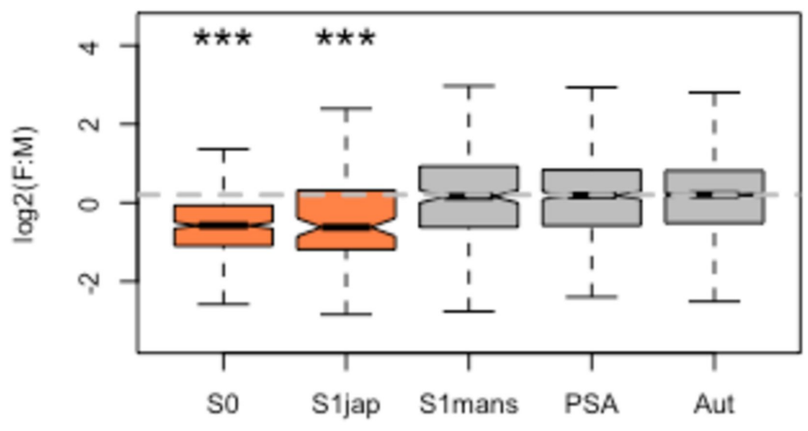


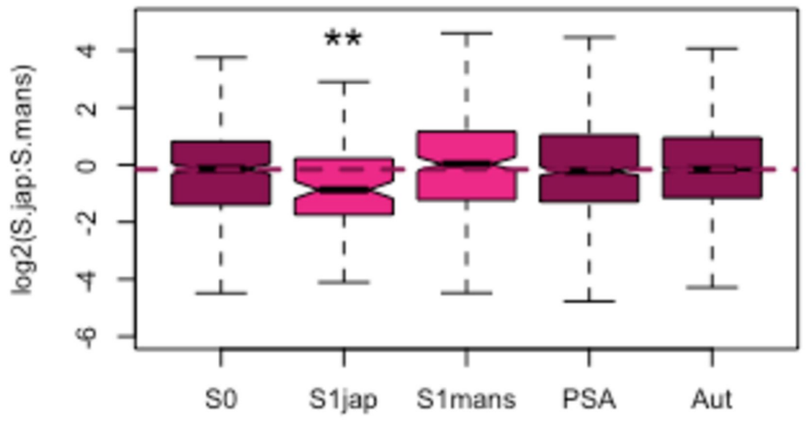


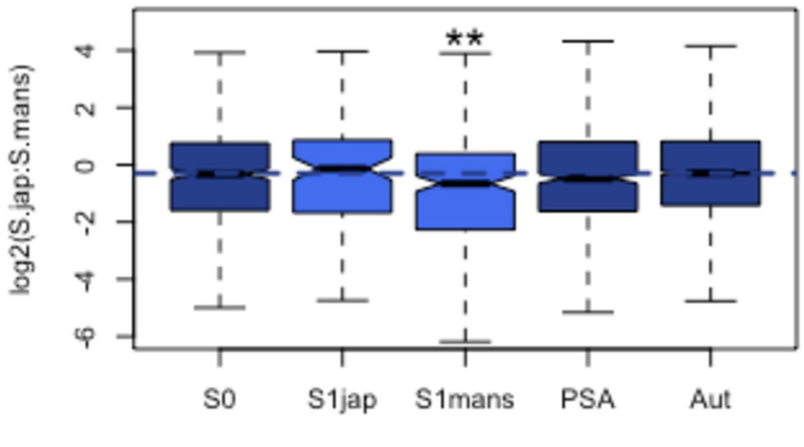


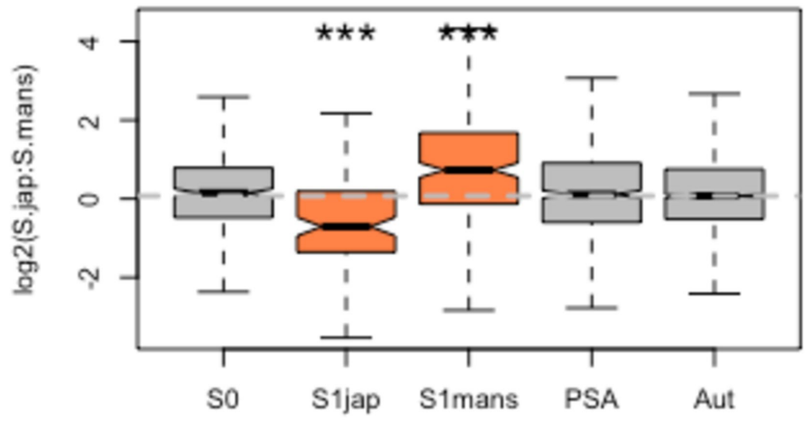




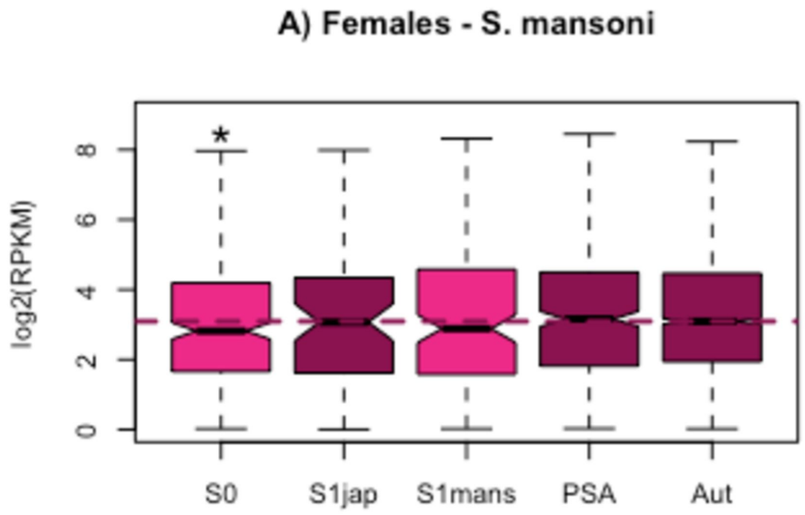


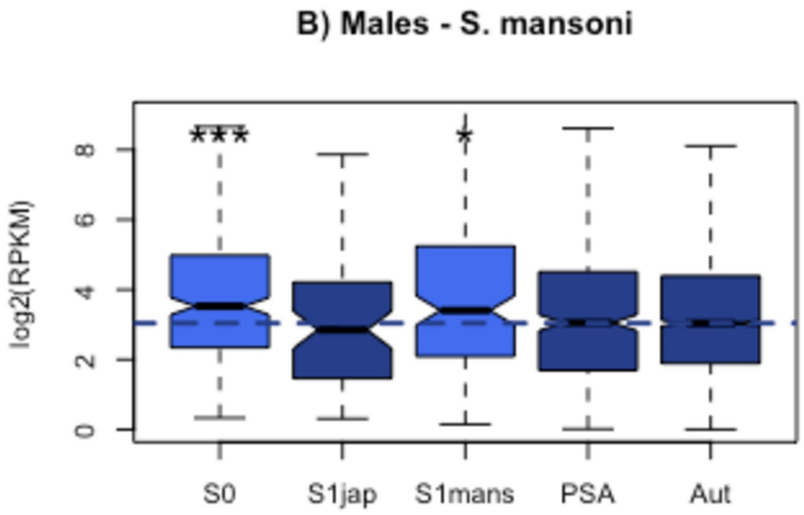


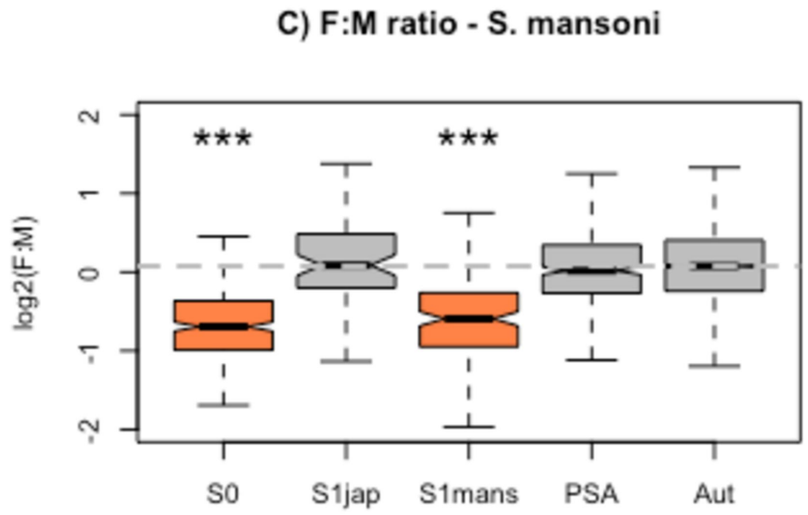


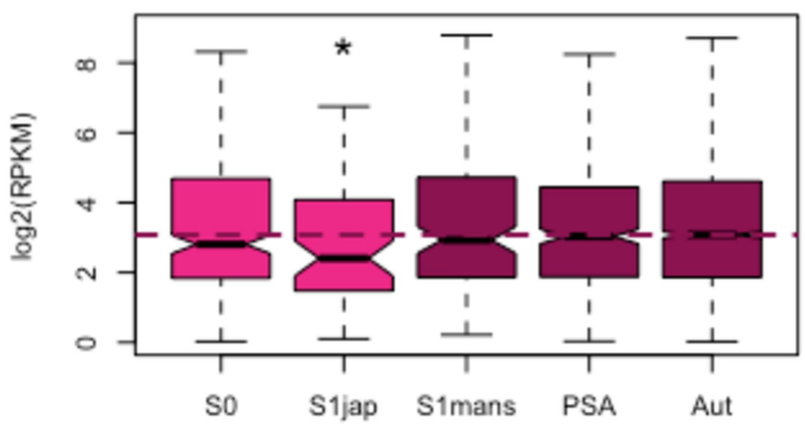


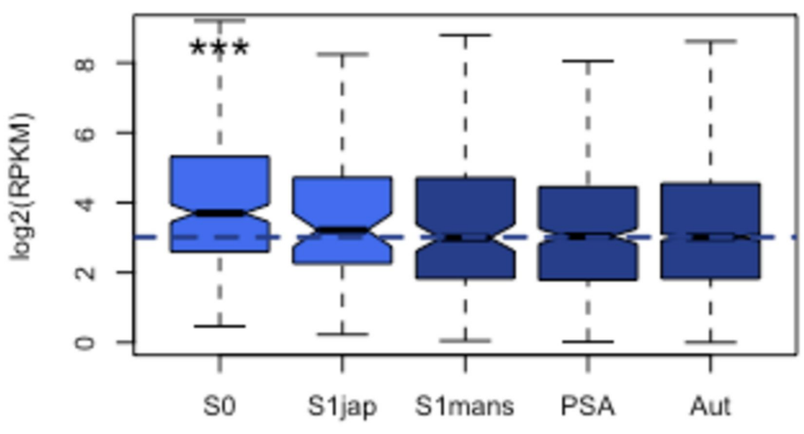
SCHISTOSOMULA - RPKM>1 - Strata Method2 (Exhaustive)

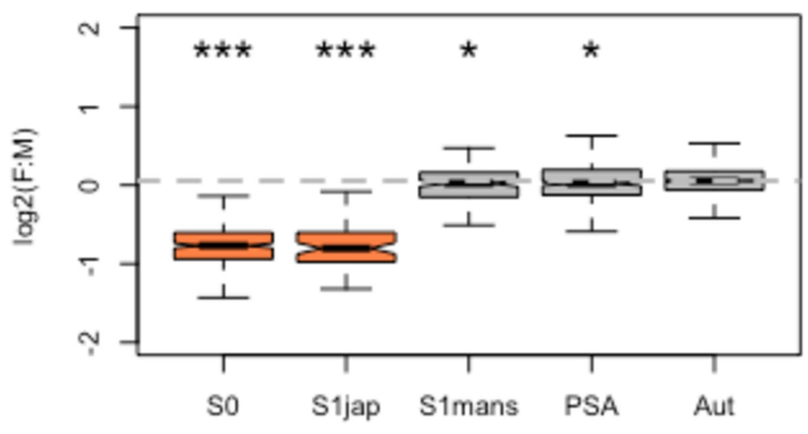


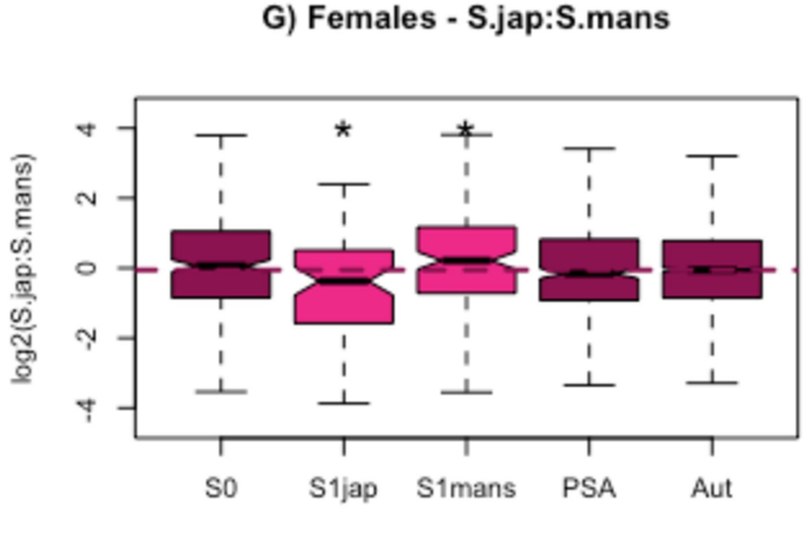


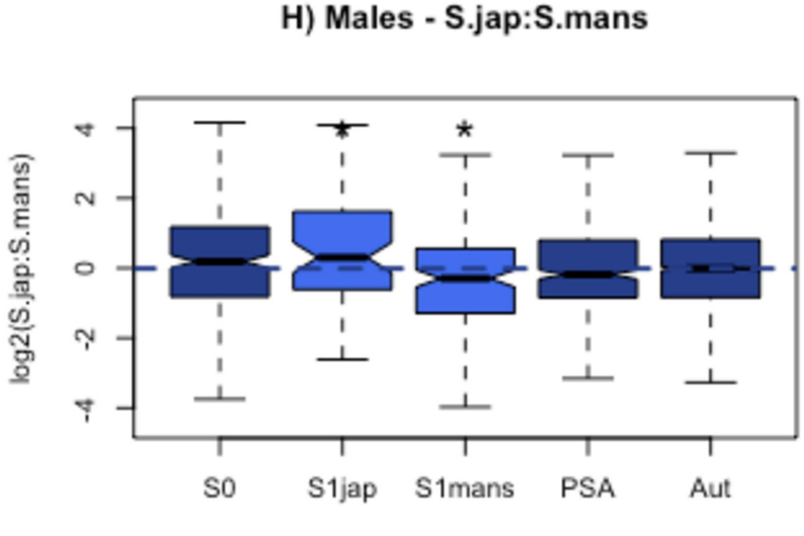


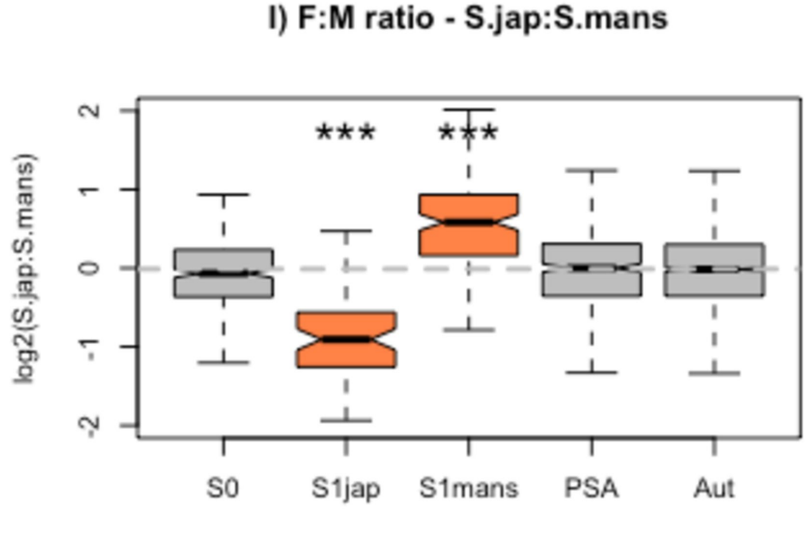




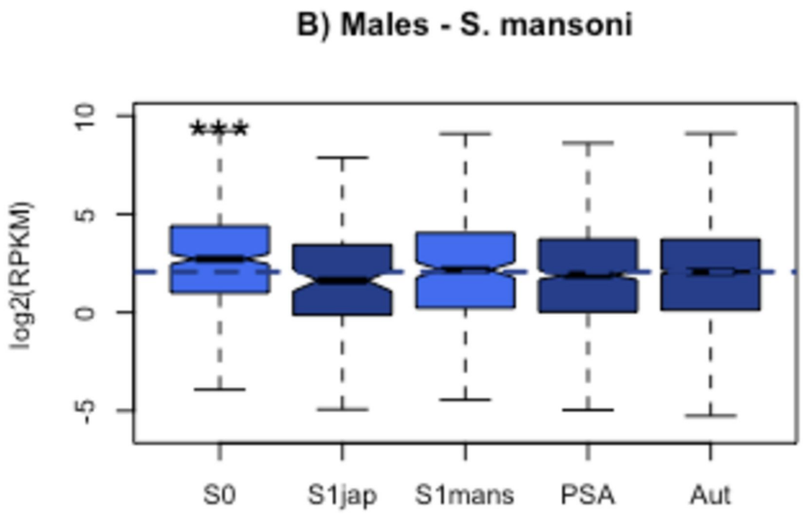


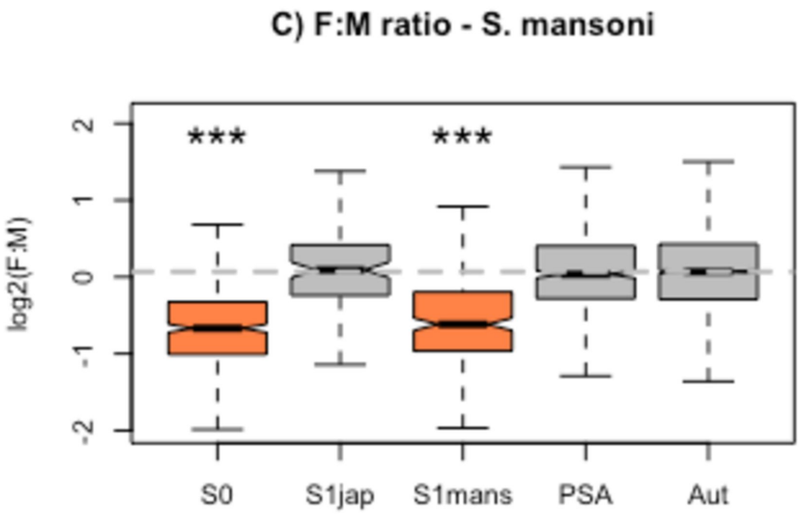


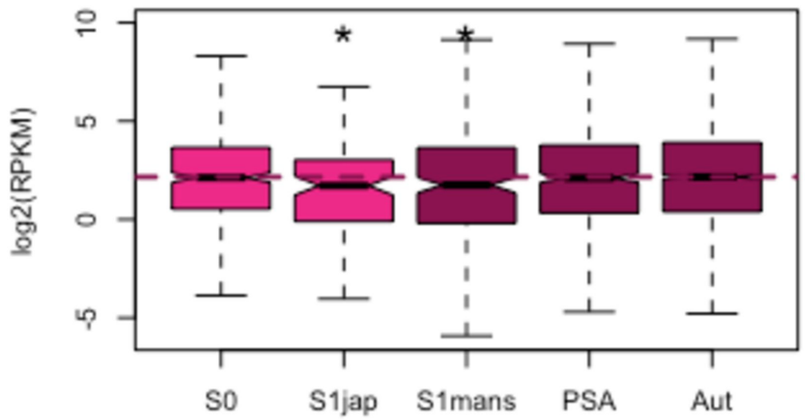


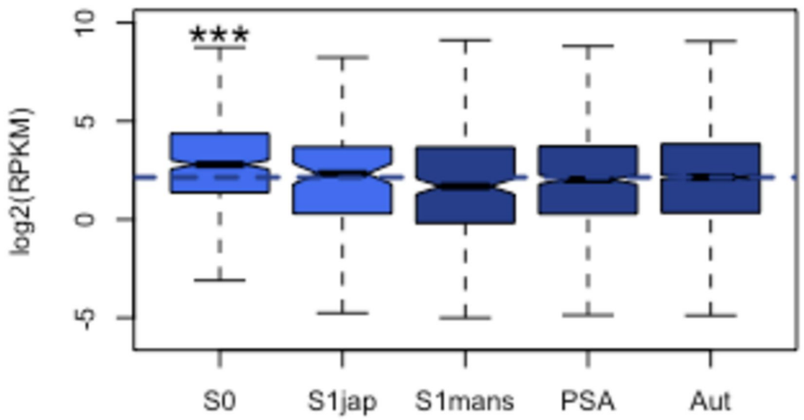


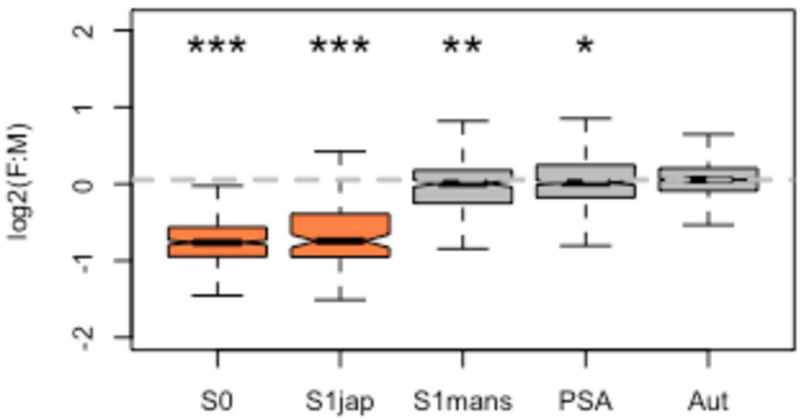
SCHISTOSOMULA - RPKM>0 - Strata Method2 (Exhaustive)

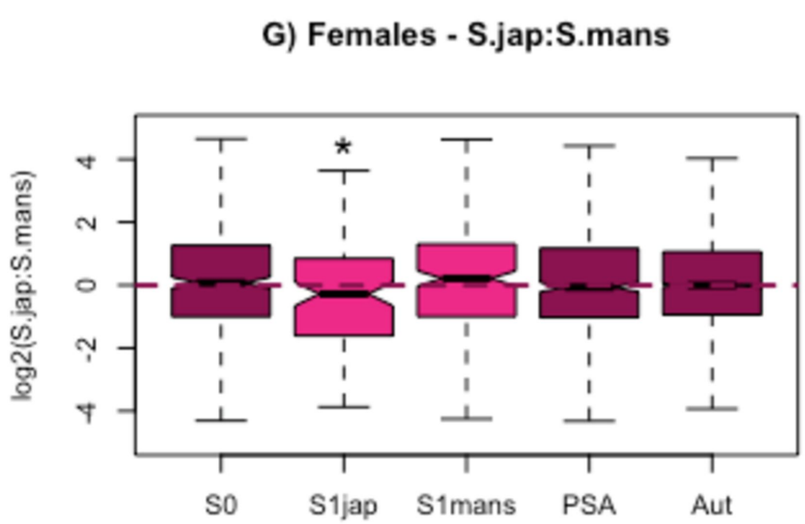


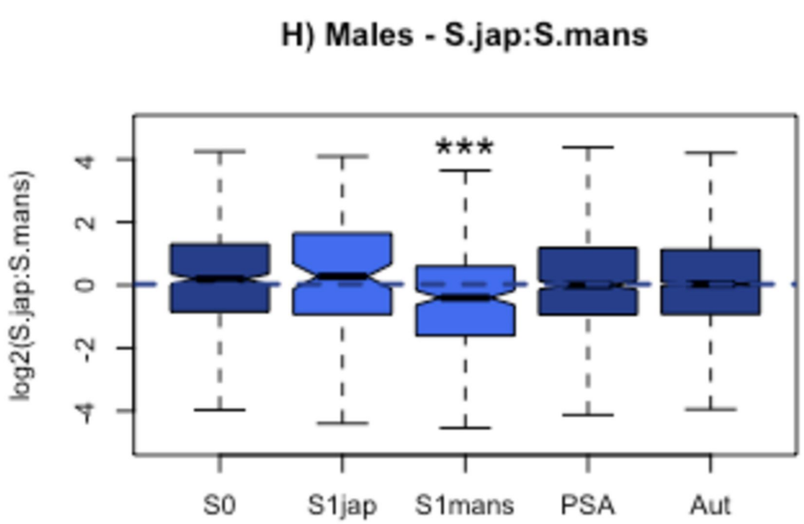


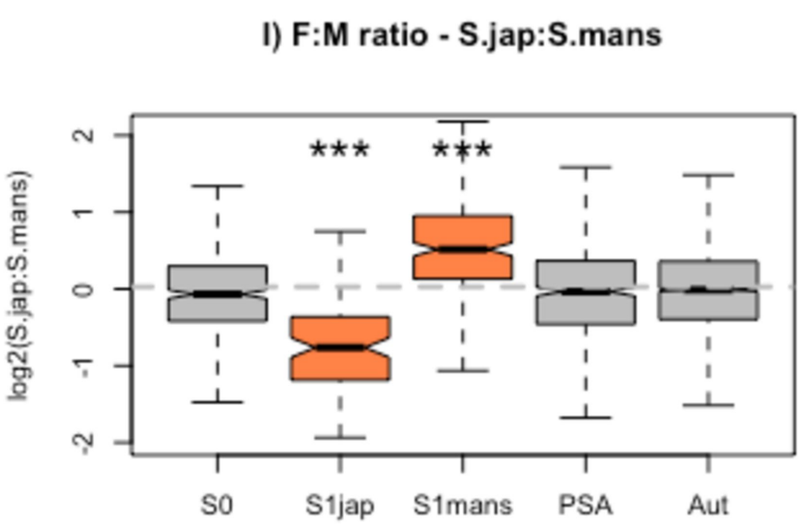




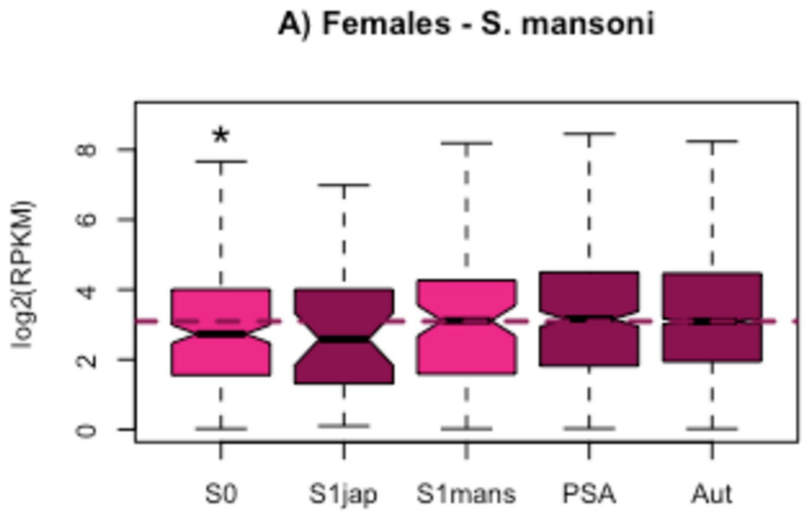


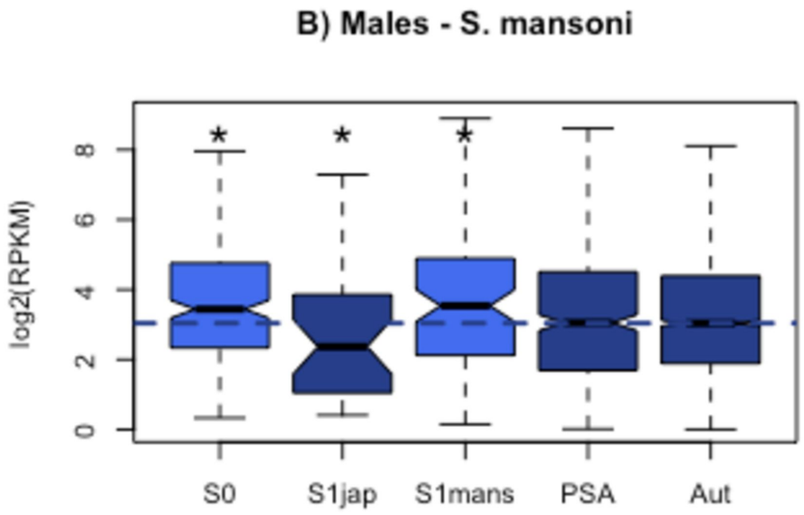


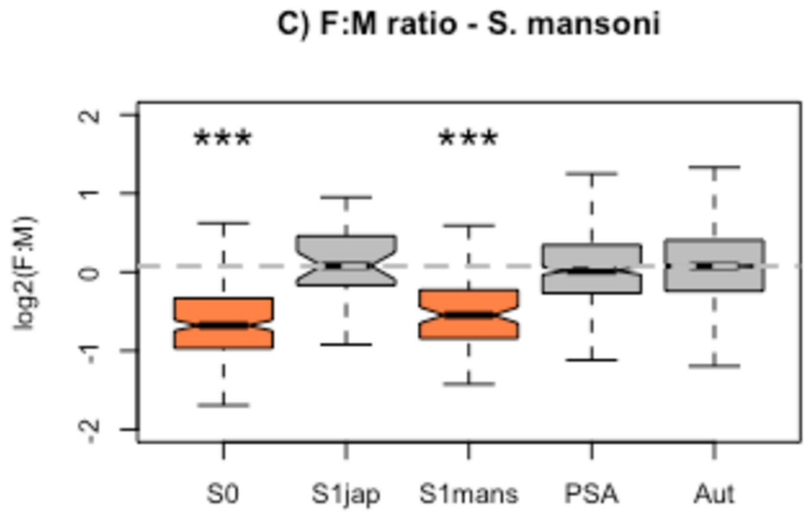


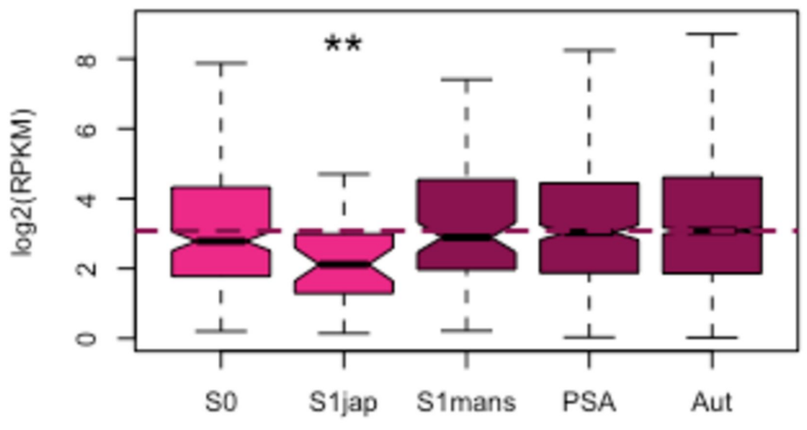


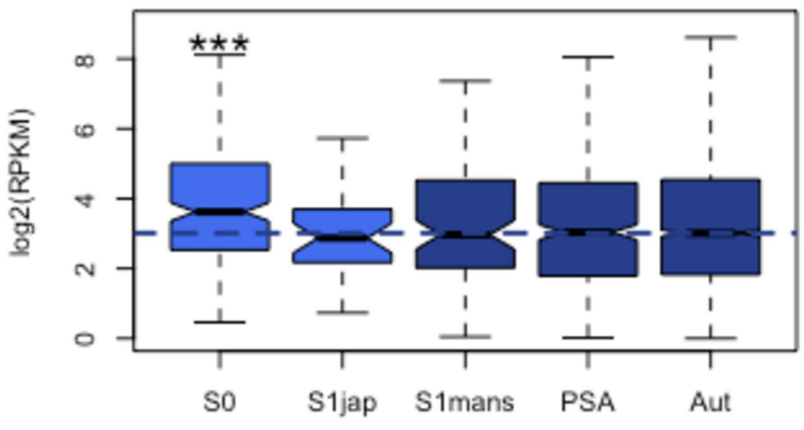
SCHISTOSOMULA - RPKM>1 - Strata Method1 (Stringent)

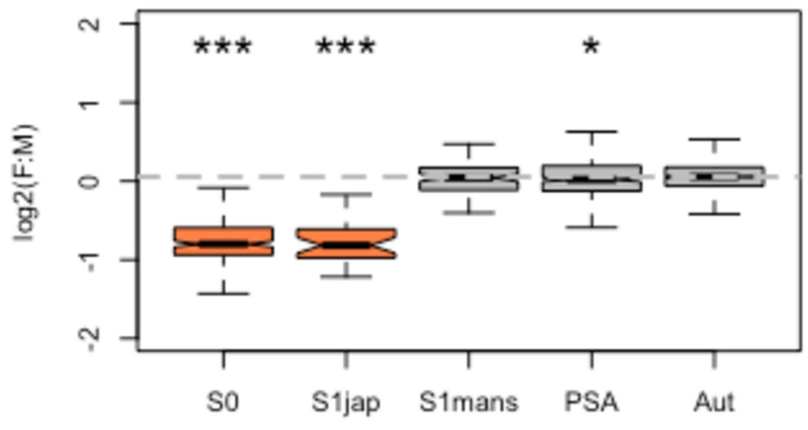


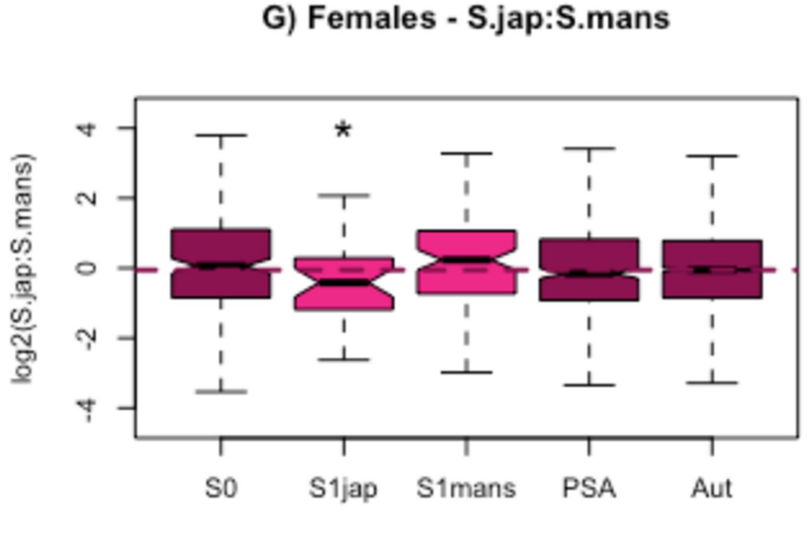


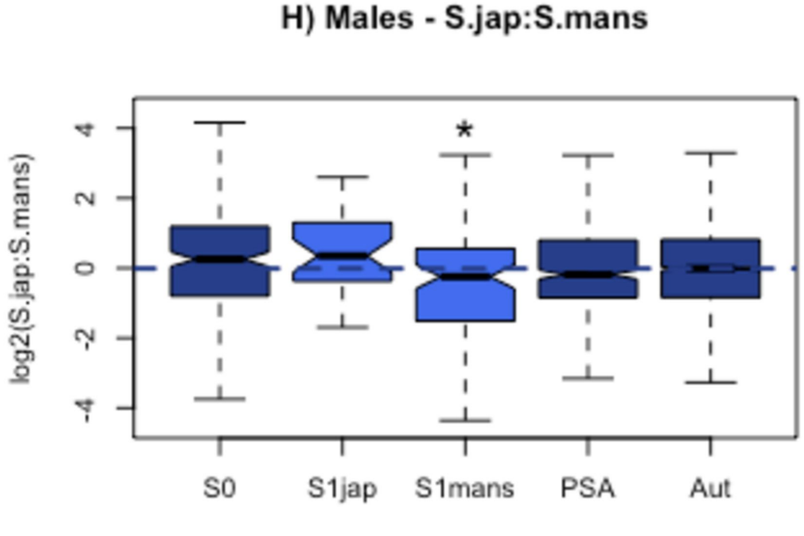


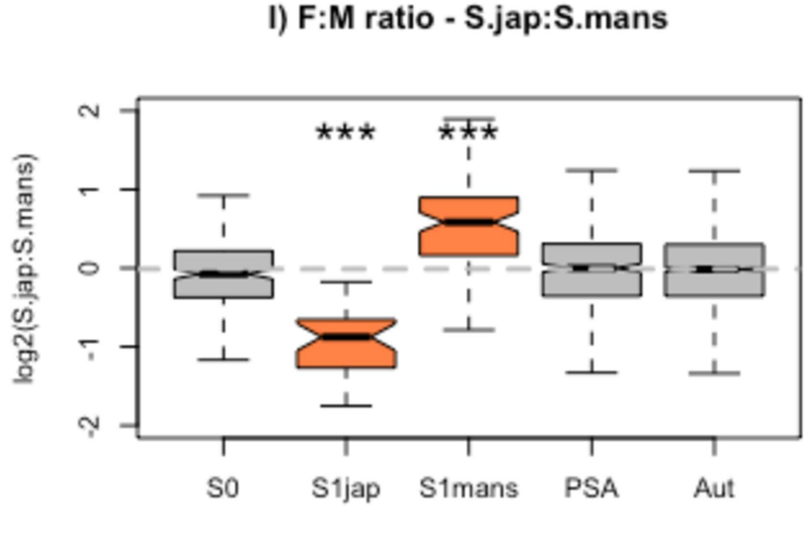




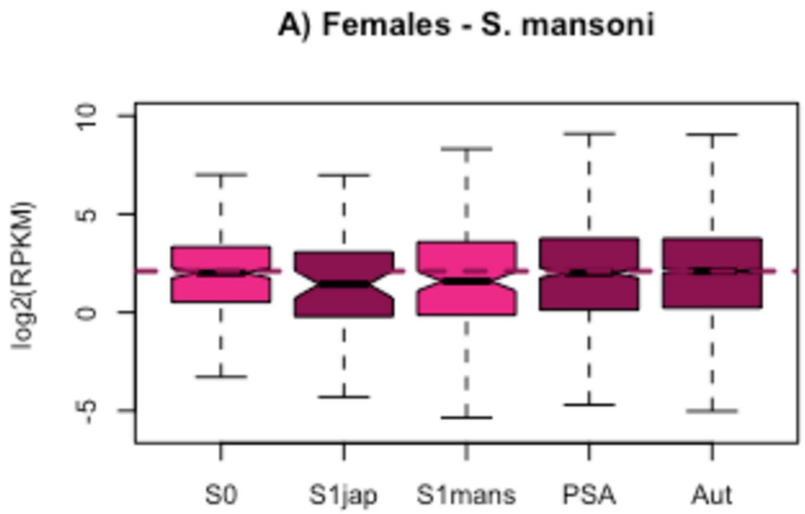


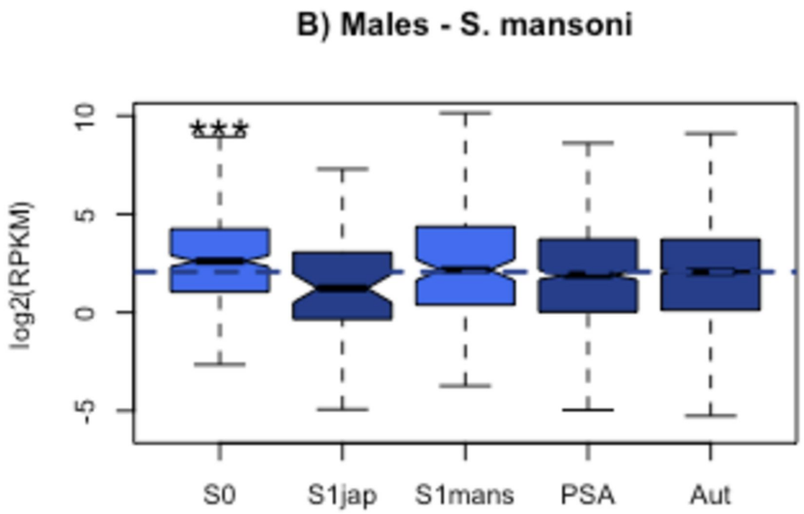


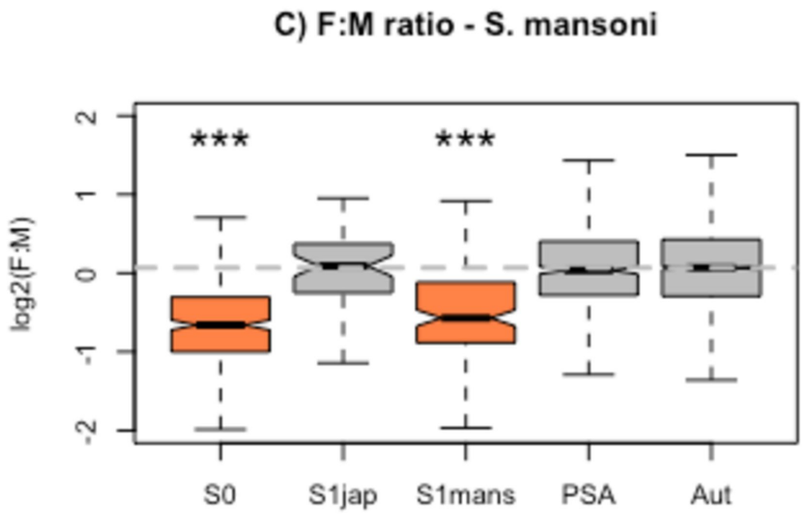


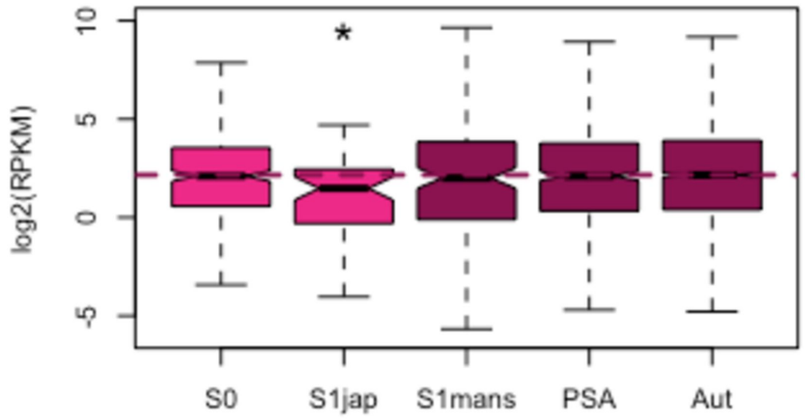


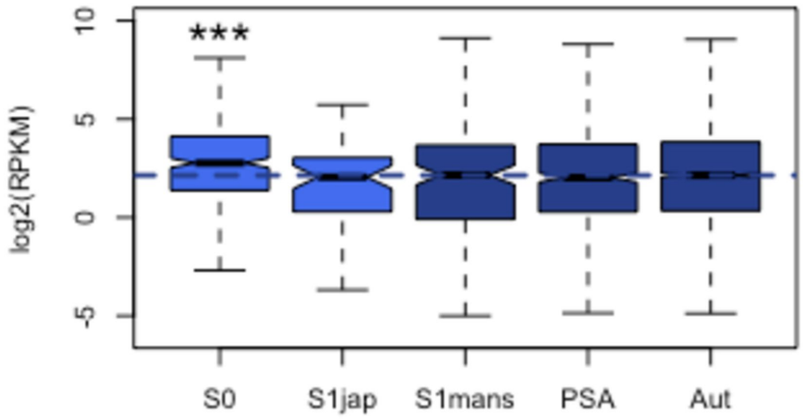
SCHISTOSOMULA - RPKM>0 - Strata Method1 (Stringent)

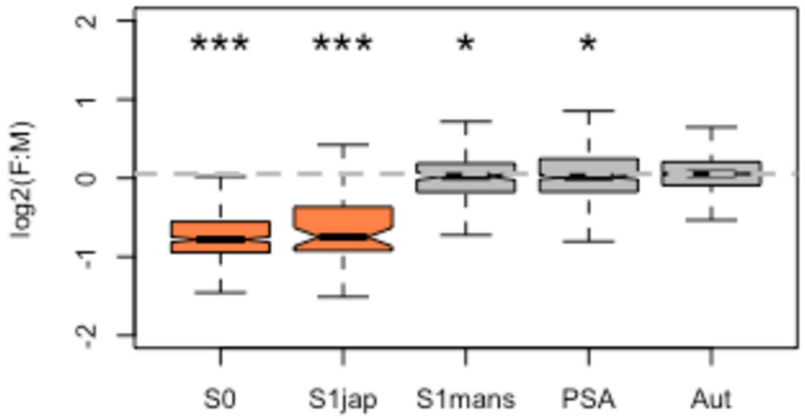


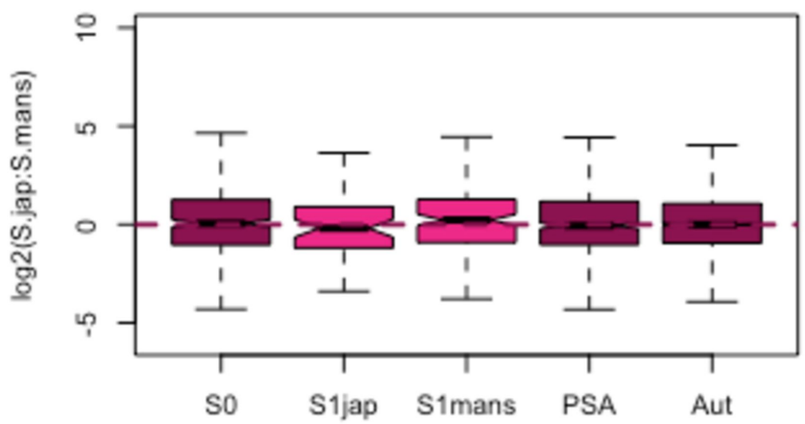


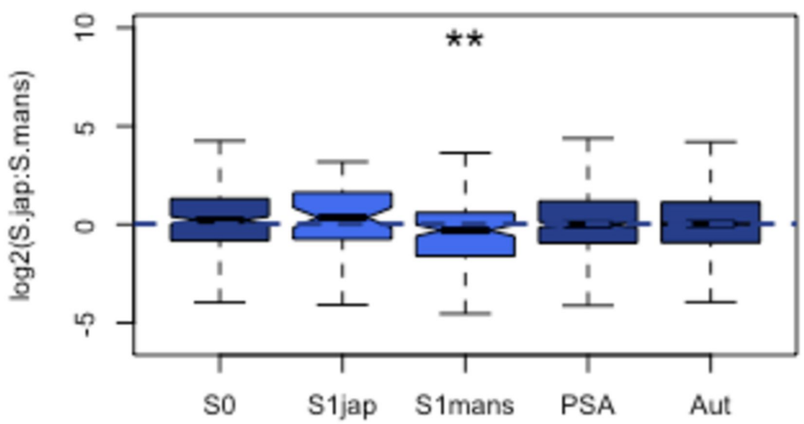


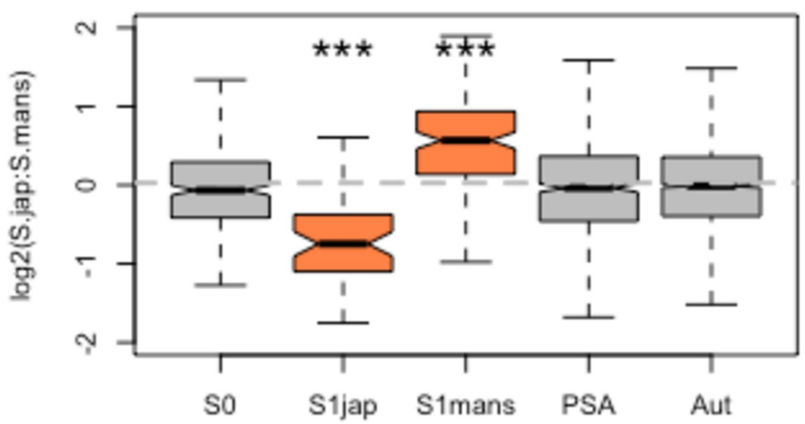




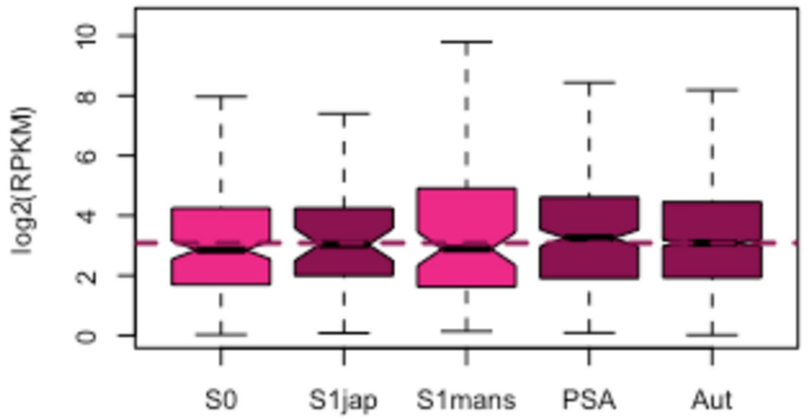




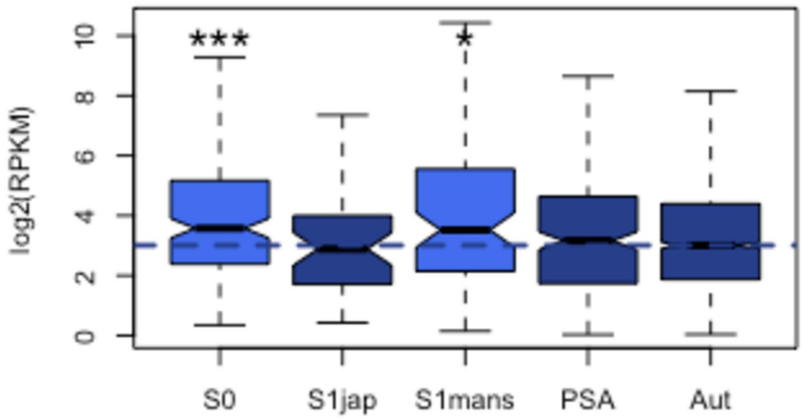




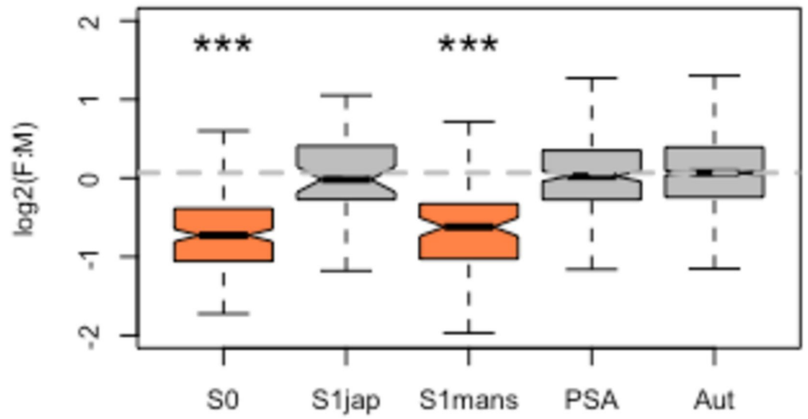
A) Females - S. mansoni



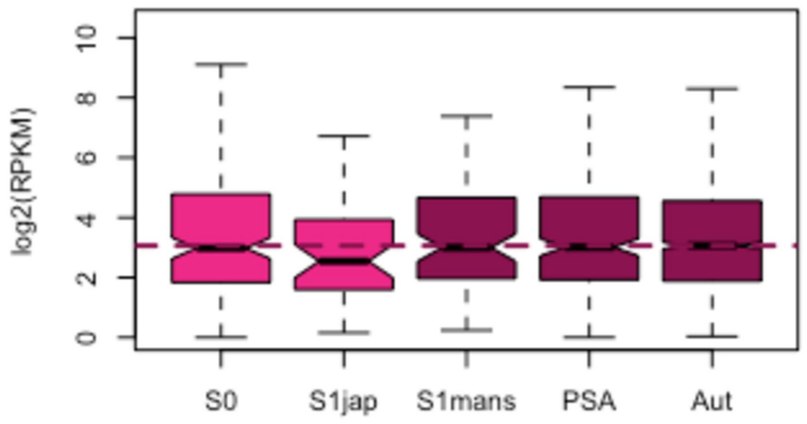
B) Males - S. mansoni



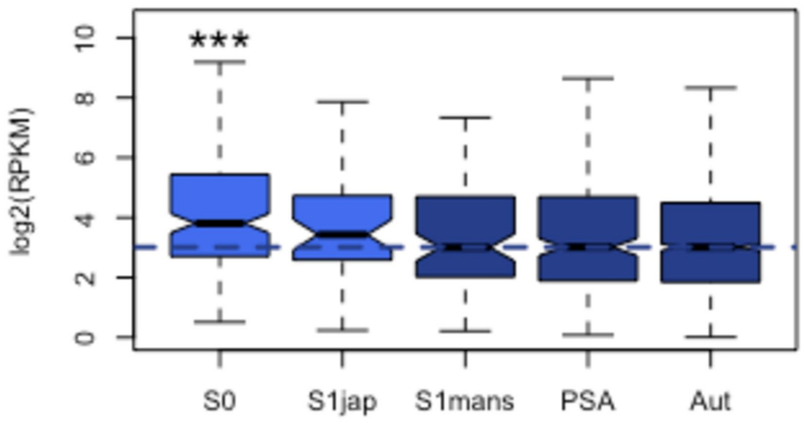
C) F:M ratio - S. mansoni



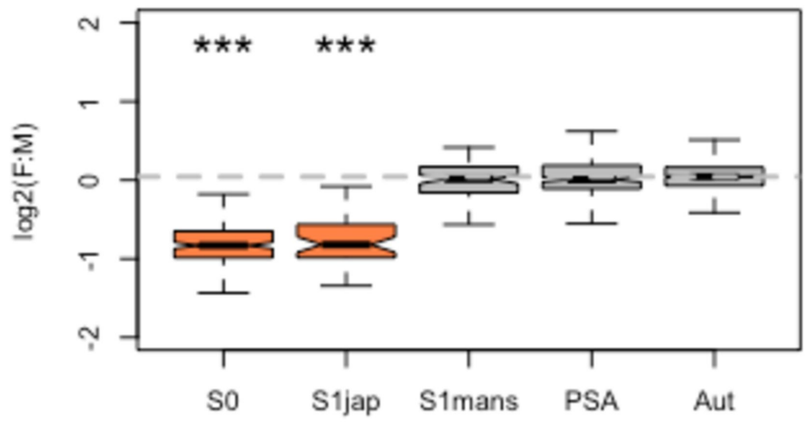
D) Females - S. japonicum



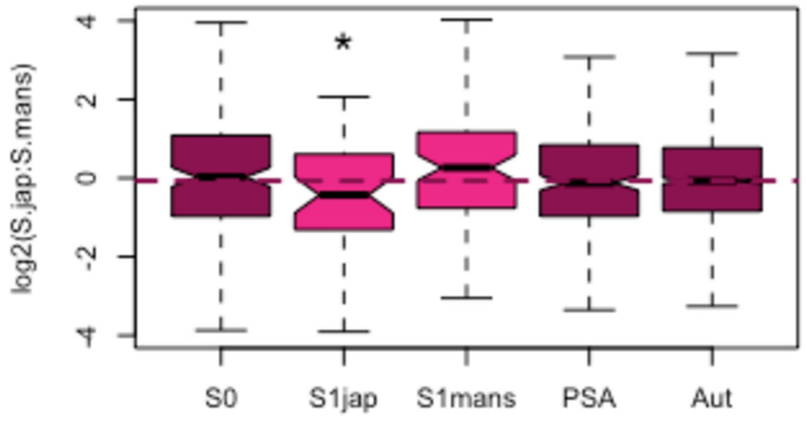
E) Males - S. japonicum



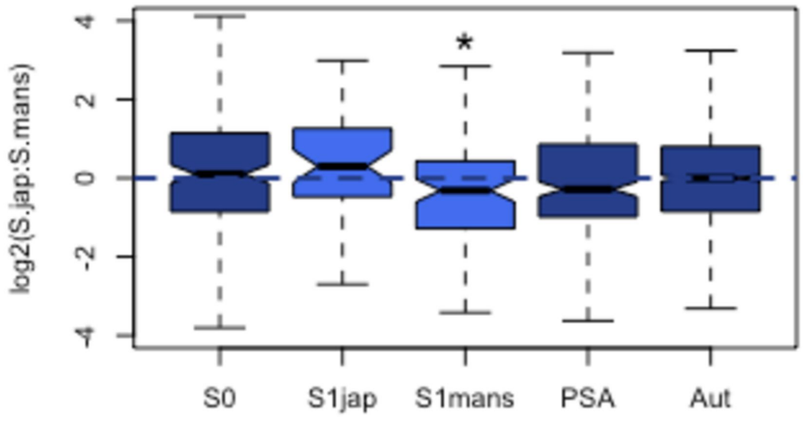
F) F:M ratio - S. japonicum



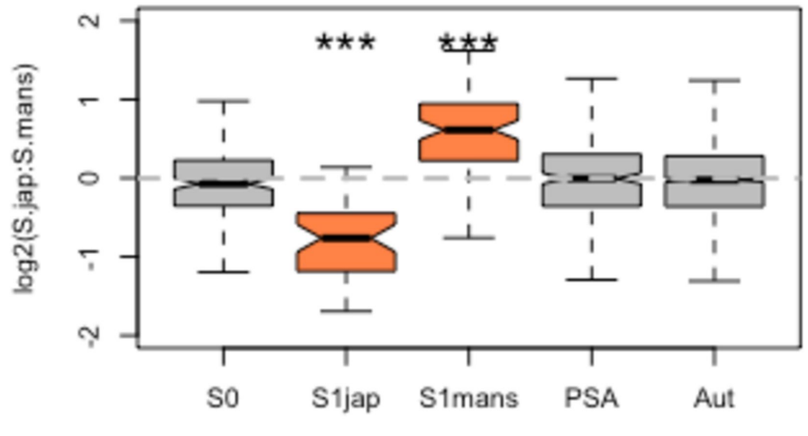
G) Females - S.jap:S.mans



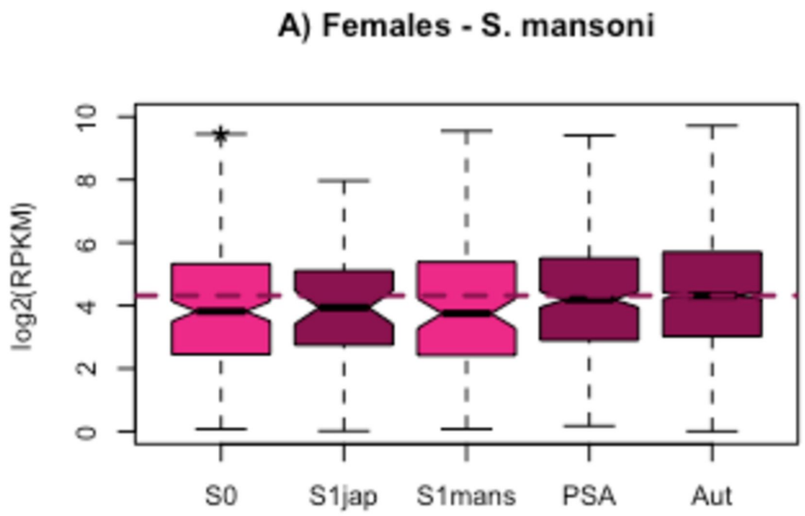
H) Males - S.jap:S.mans

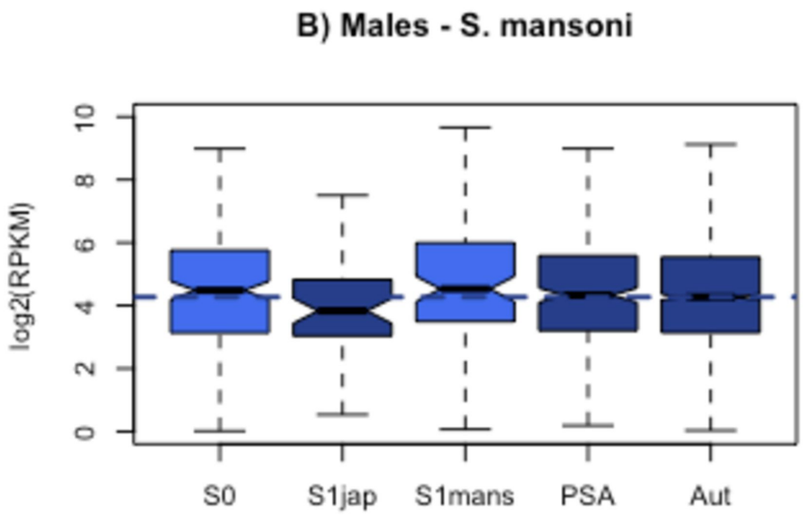


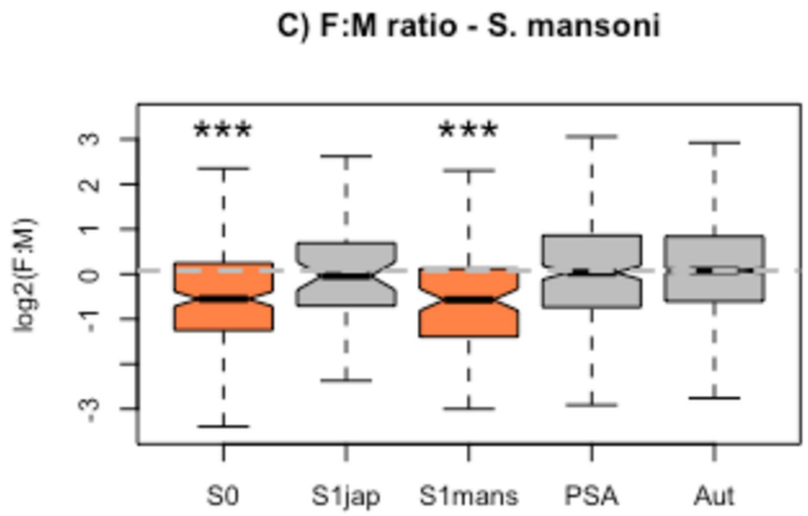
I) F:M ratio - S.jap:S.mans

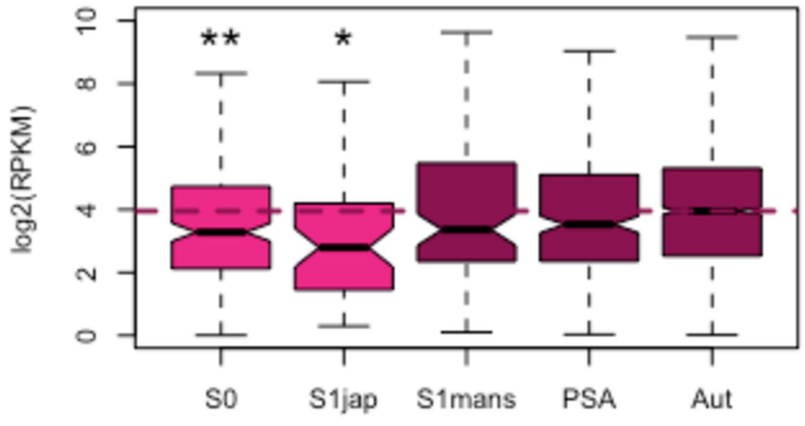


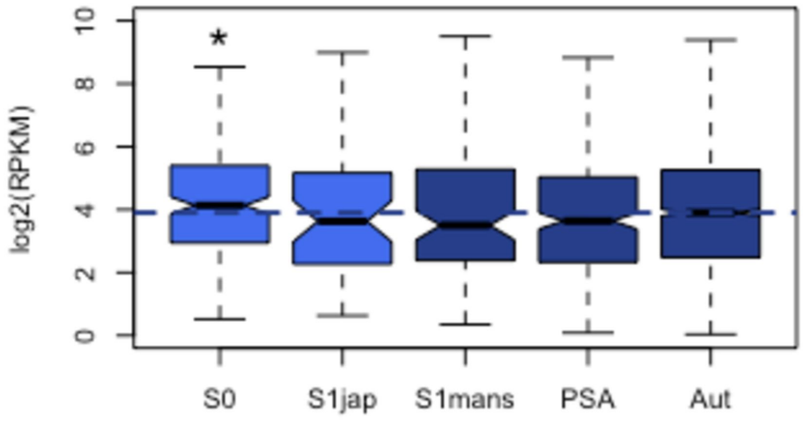
SCHISTOSOMULA - TPM>1 - Strata Method2 (Exhaustive)

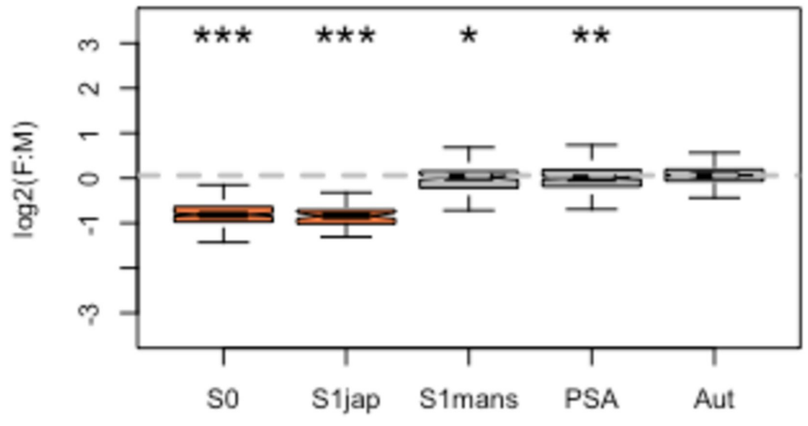


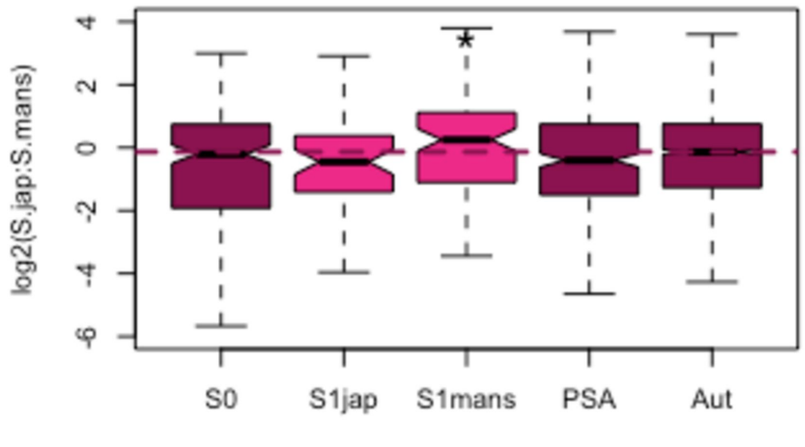


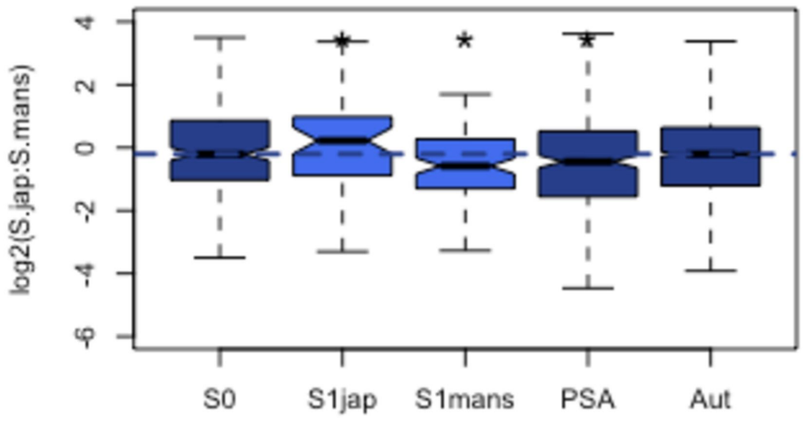


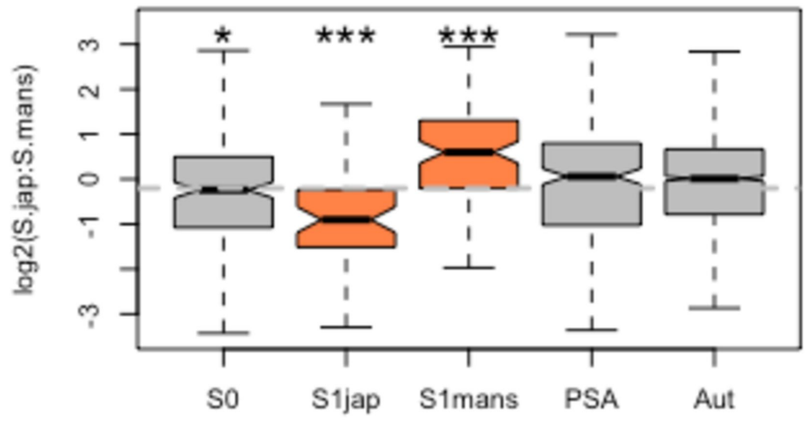




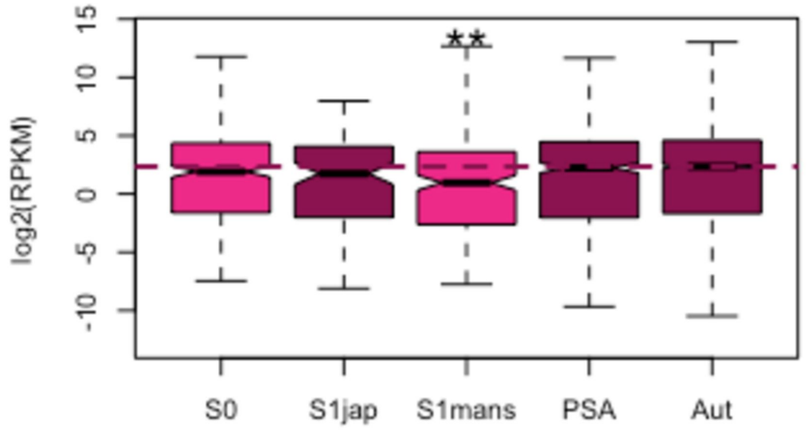




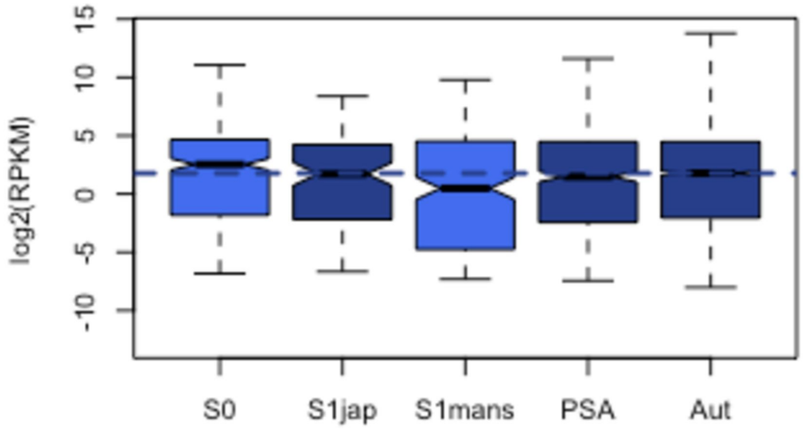




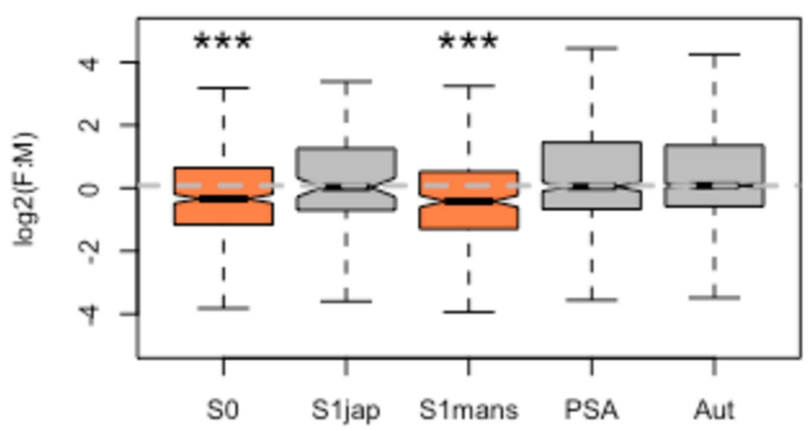
A) Females - S. mansoni



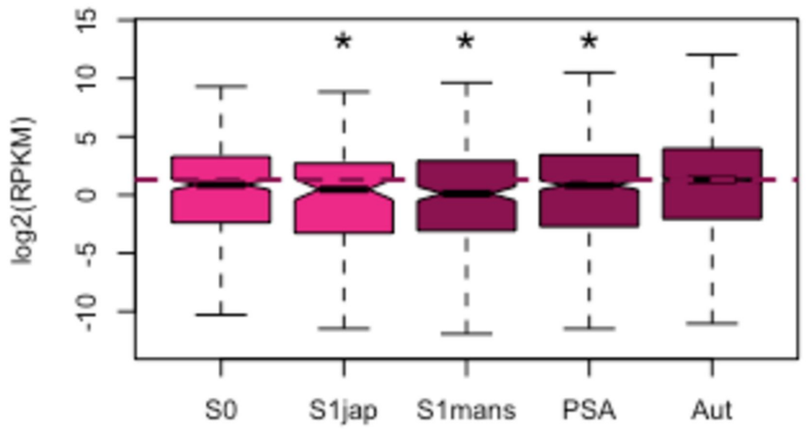
B) Males - S. mansoni



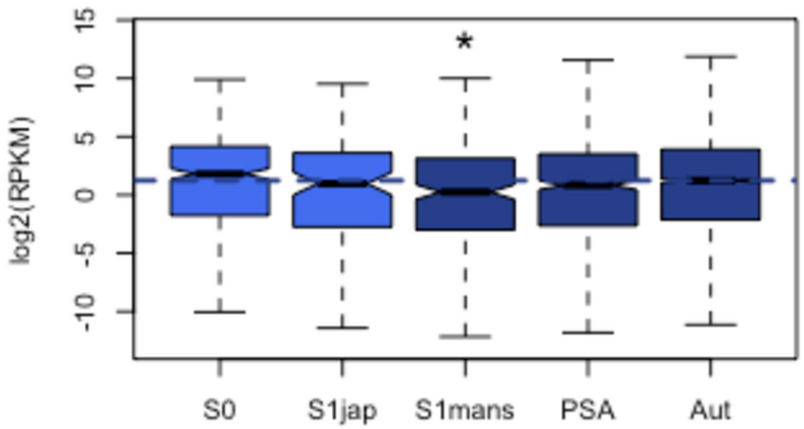
C) F:M ratio - S. mansoni



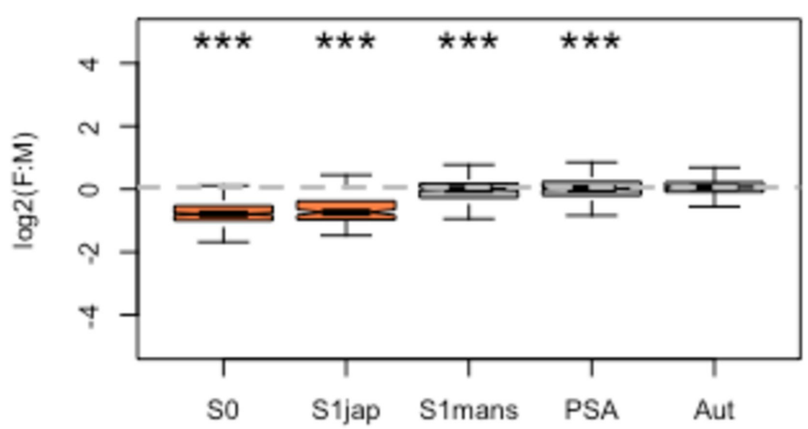
D) Females - S. japonicum



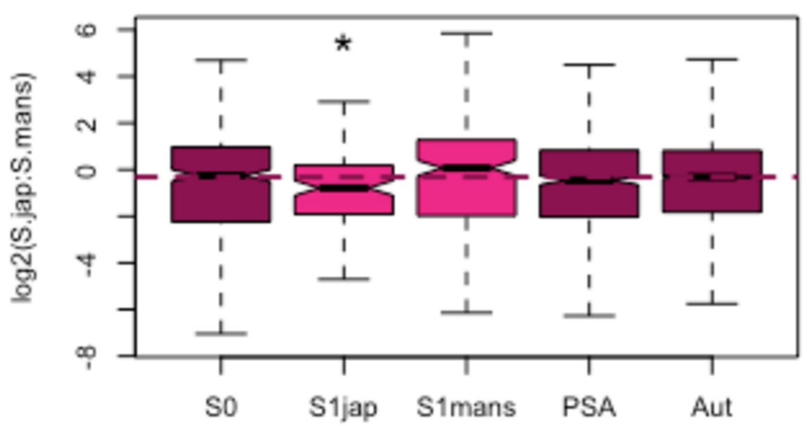
E) Males - S. japonicum



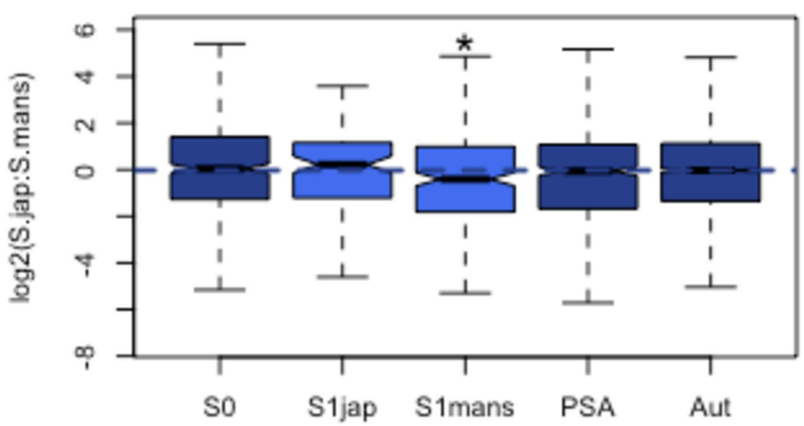
F) F:M ratio - S. japonicum



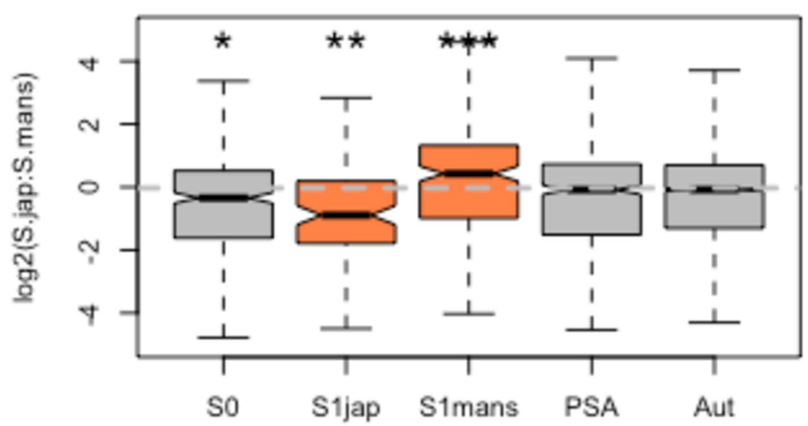
G) Females - S.jap:S.mans

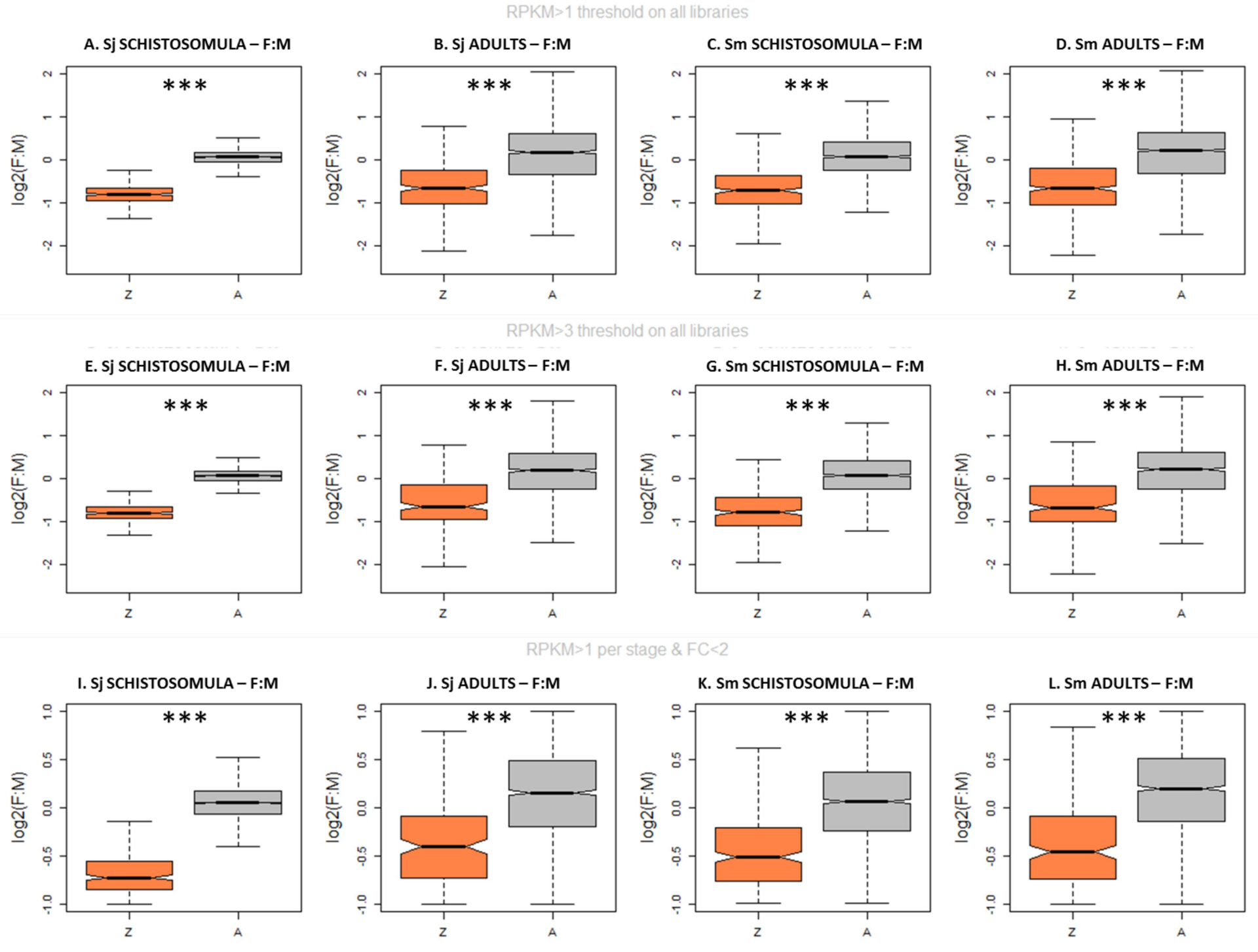


H) Males - S.jap:S.mans

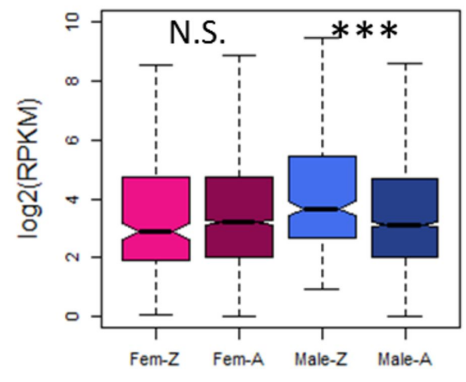


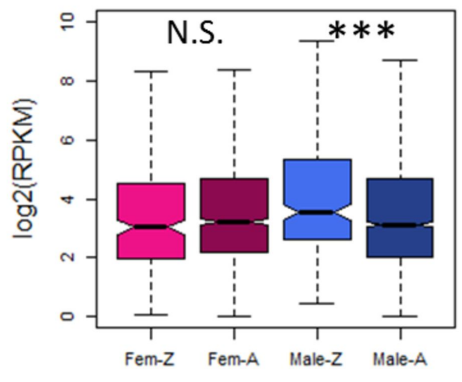
I) F:M ratio - S.jap:S.mans

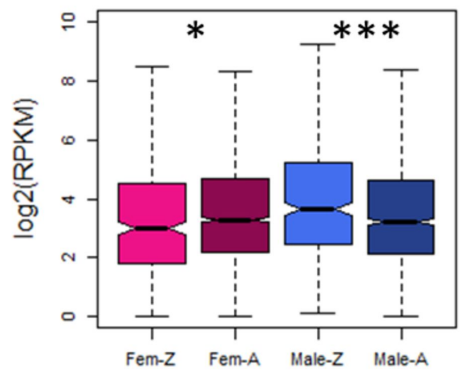


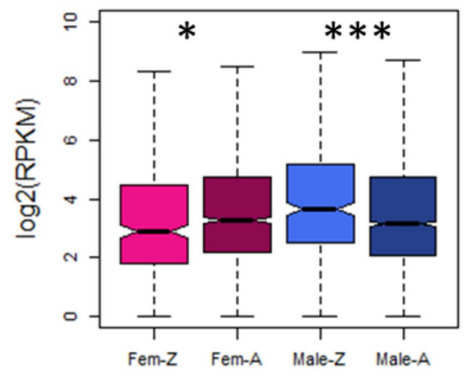

RPKM>1 threshold on all libraries
A. Sj SCHISTOSOMULA – F:M
B. Sj ADULTS – F:M
C. Sm SCHISTOSOMULA – F:M
D. Sm ADULTS – F:M
RPKM>3 threshold on all libraries
E. Sj SCHISTOSOMULA – F:M
F. Sj ADULTS – F:M
G. Sm SCHISTOSOMULA – F:M
H. Sm ADULTS – F:M
RPKM>1 per stage & FC<2
I. Sj SCHISTOSOMULA – F:M
J. Sj ADULTS – F:M
K. Sm SCHISTOSOMULA – F:M
L. Sm ADULTS – F:M
log2(F:M)
***
Z
A

RPKM>1 threshold on all libraries

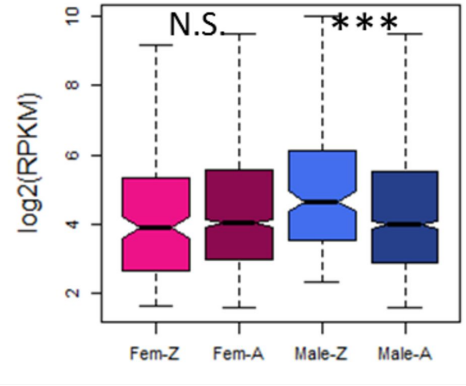


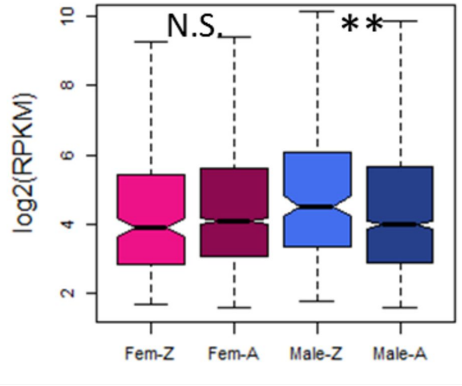


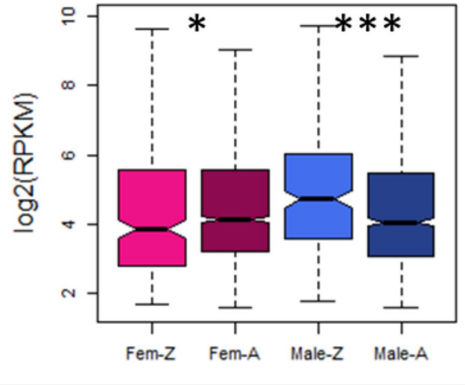


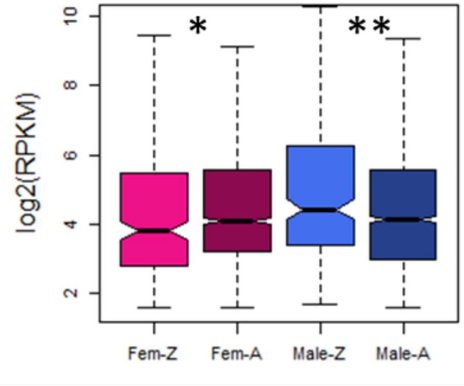


RPKM>3 threshold on all libraries

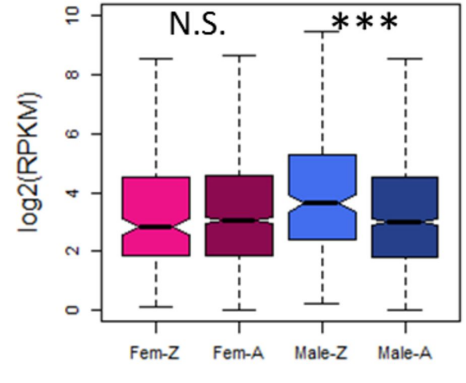


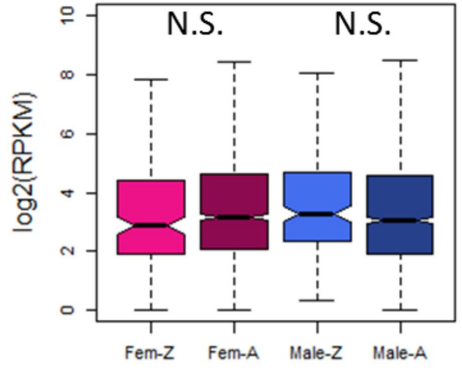


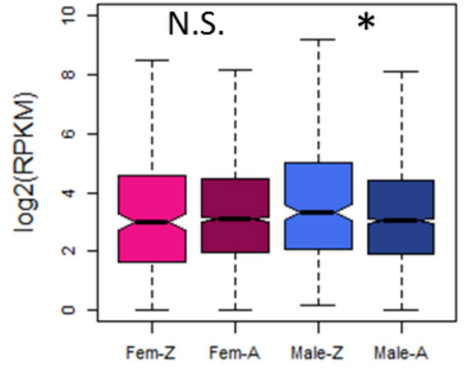


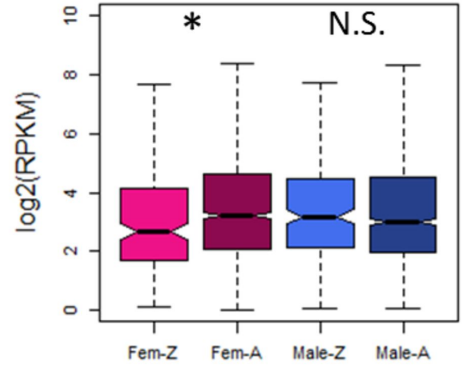


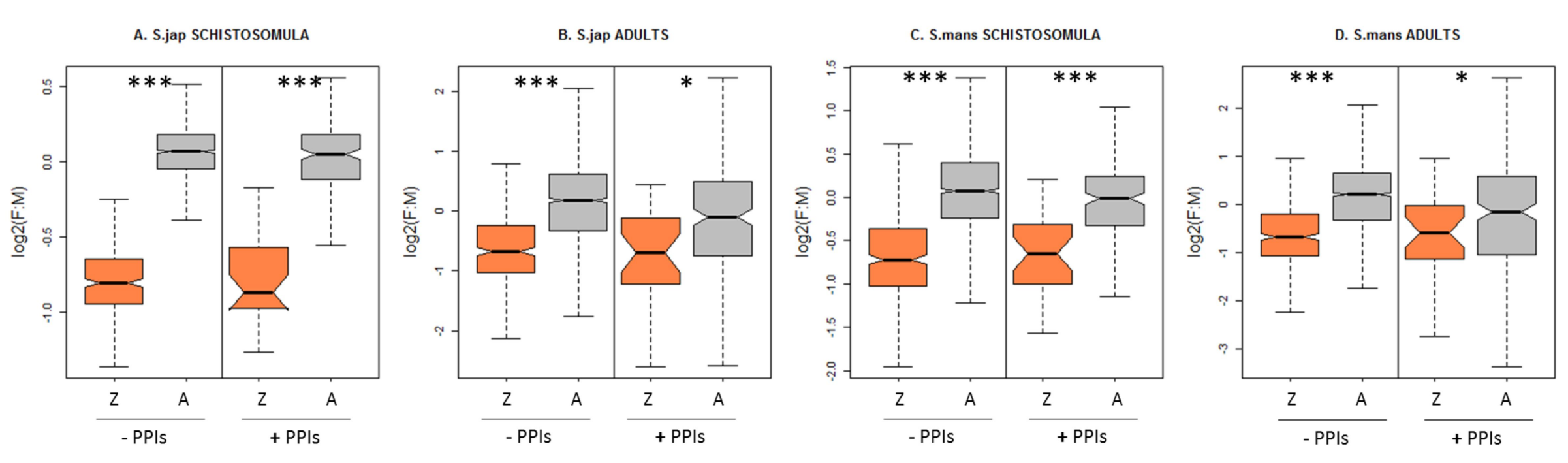
RPKM>1 per stage & FC<2

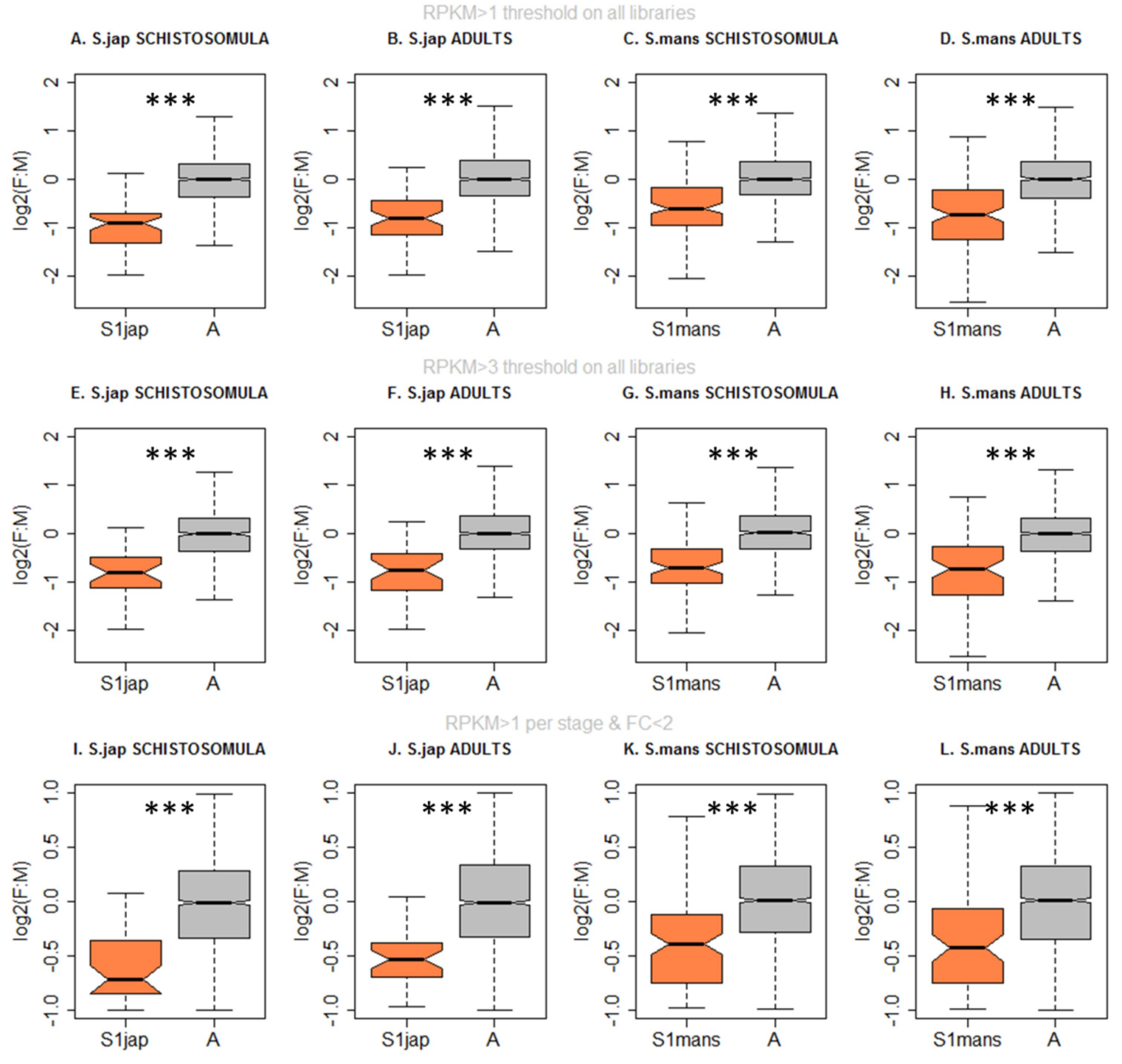


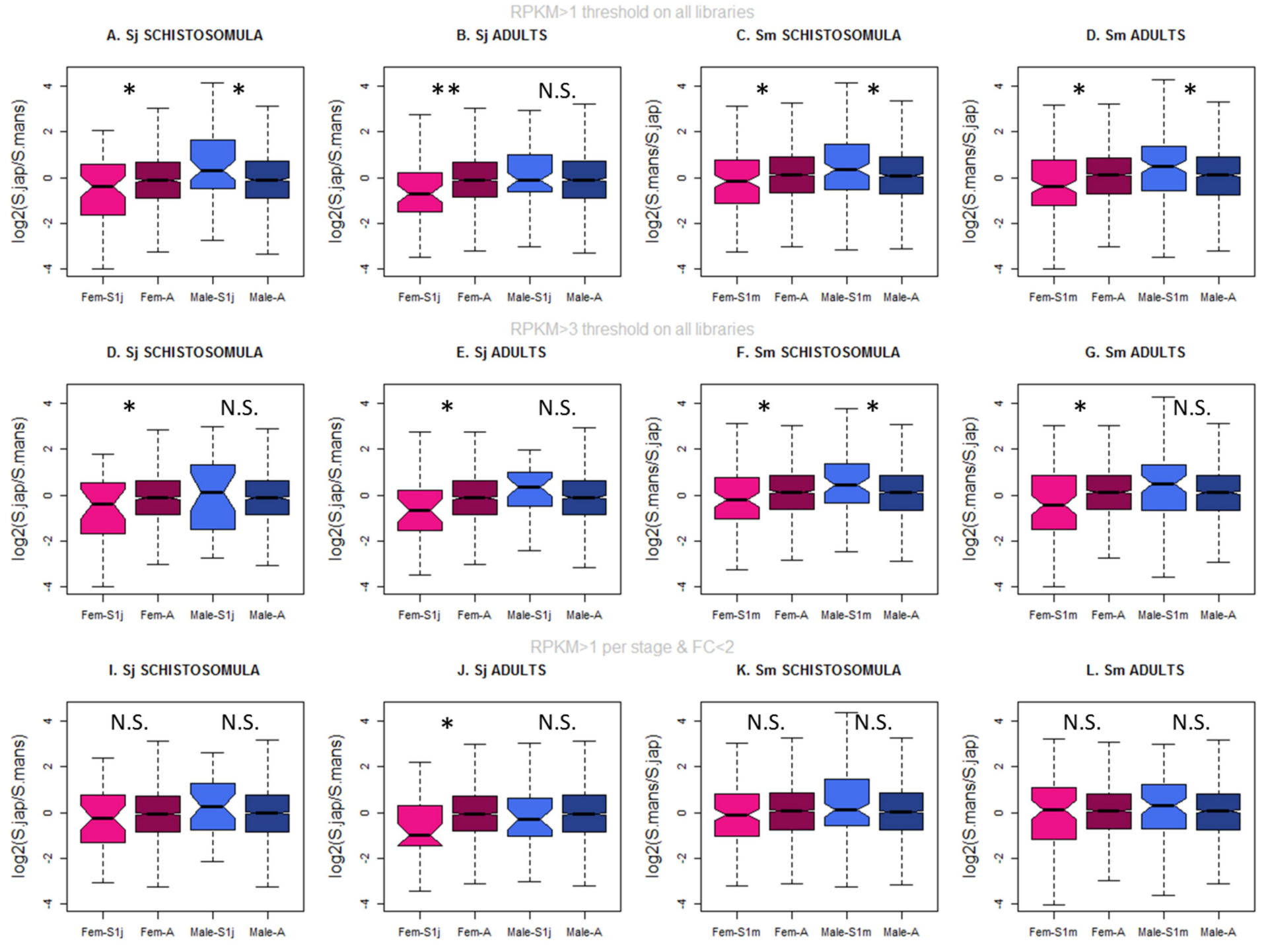


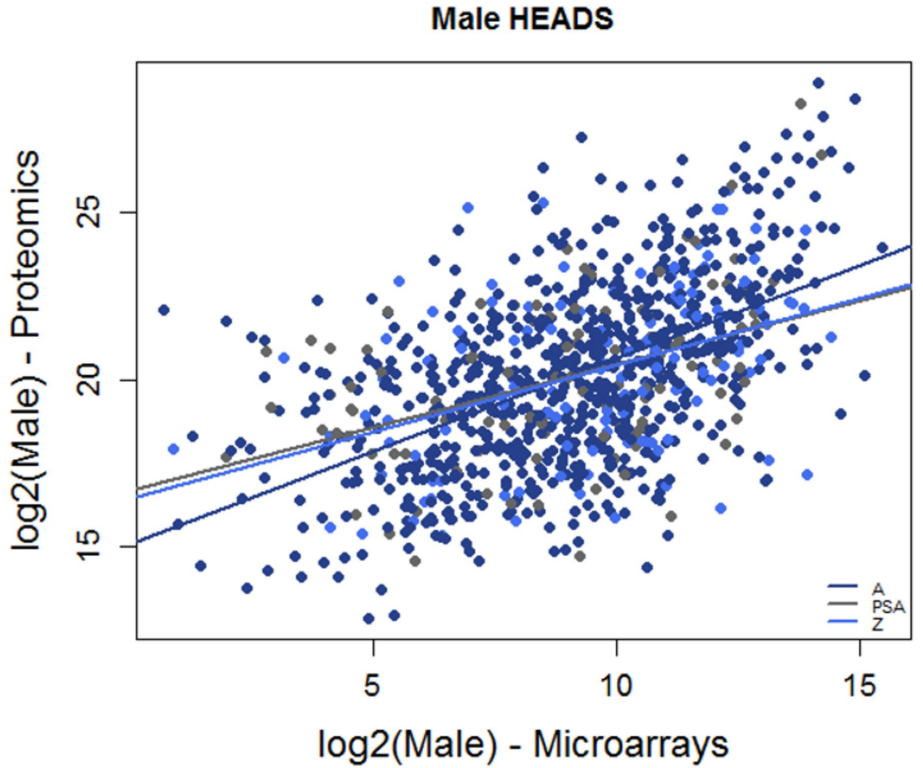





A. S.jap SCHISTOSOMULA
B. S.jap ADULTS
C. S.mans SCHISTOSOMULA
D. S.mans ADULTS
log2(F:M)
***
*
Z
A
- PPIs
+ PPIs

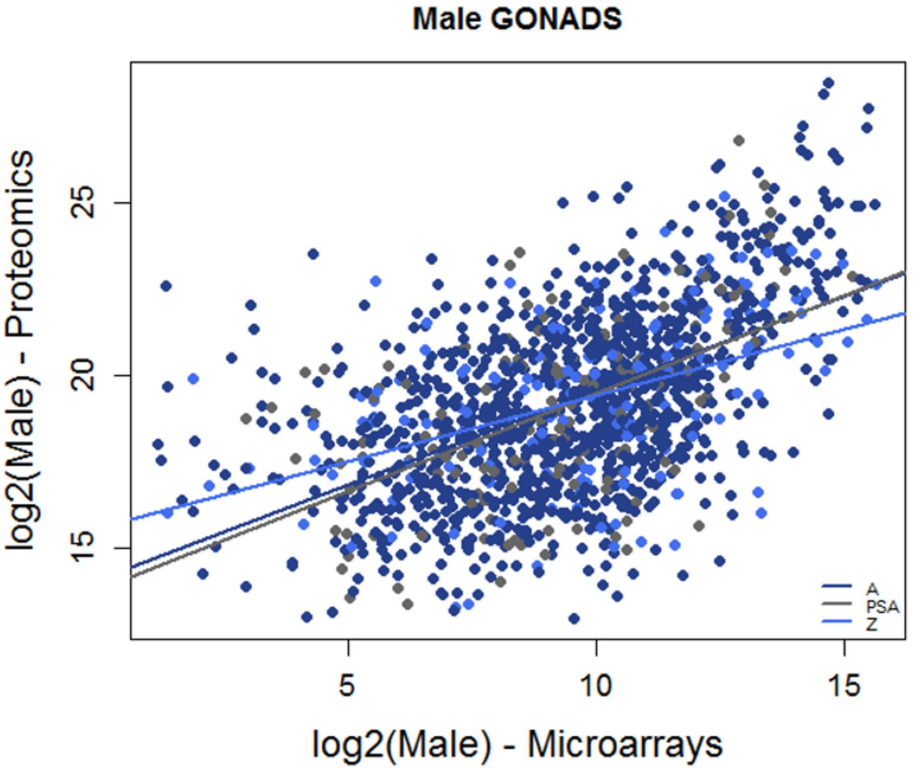

RPKM>1 threshold on all libraries
A. S.jap SCHISTOSOMULA
B. S.jap ADULTS
C. S.mans SCHISTOSOMULA
D. S.mans ADULTS
RPKM>3 threshold on all libraries
E. S.jap SCHISTOSOMULA
F. S.jap ADULTS
G. S.mans SCHISTOSOMULA
H. S.mans ADULTS
RPKM>1 per stage & FC<2
I. S.jap SCHISTOSOMULA
J. S.jap ADULTS
K. S.mans SCHISTOSOMULA
L. S.mans ADULTS
log2(F:M)
***
S1jap
S1mans
A


RPKM>1 threshold on all libraries
A. Sj SCHISTOSOMULA
B. Sj ADULTS
C. Sm SCHISTOSOMULA
D. Sm ADULTS
log2(S.jap/S.mans)
log2(S.mans/S.jap)
*
**
N.S.
Fem-S1j
Fem-A
Male-S1j
Male-A
Fem-S1m
Male-S1m
RPKM>3 threshold on all libraries
D. Sj SCHISTOSOMULA
E. Sj ADULTS
F. Sm SCHISTOSOMULA
G. Sm ADULTS
RPKM>1 per stage & FC<2
I. Sj SCHISTOSOMULA
J. Sj ADULTS
K. Sm SCHISTOSOMULA
L. Sm ADULTS



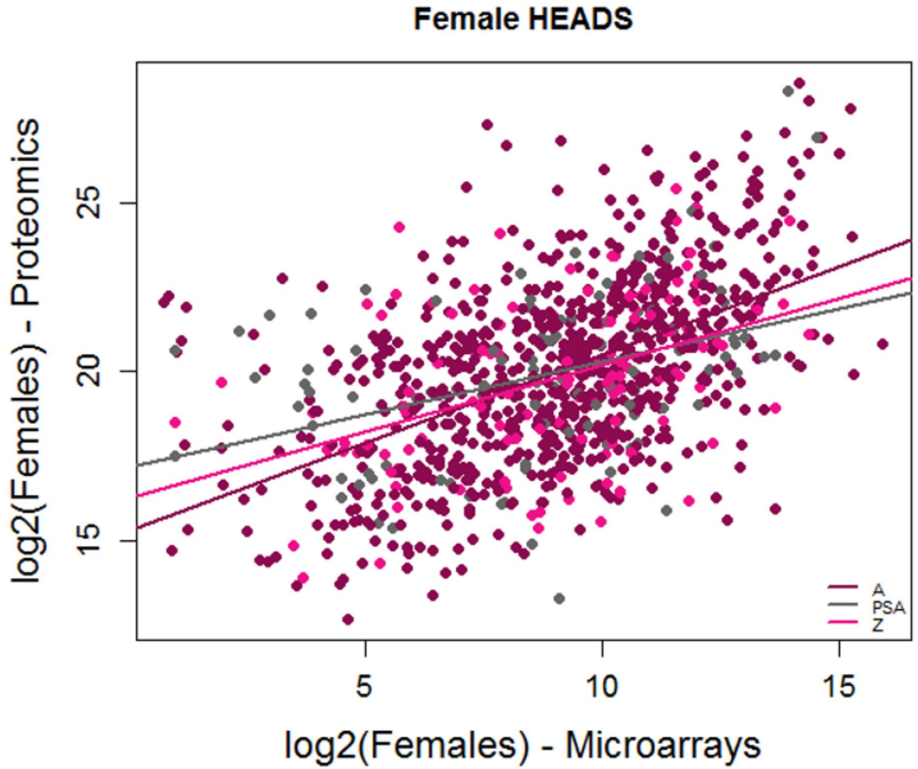
HEADS
A = 738 genes
PSA = 108 genes
Z = 123 genes

R(A) = 0.54***
R(PSA) =0.44***
R(Z) =0.43***



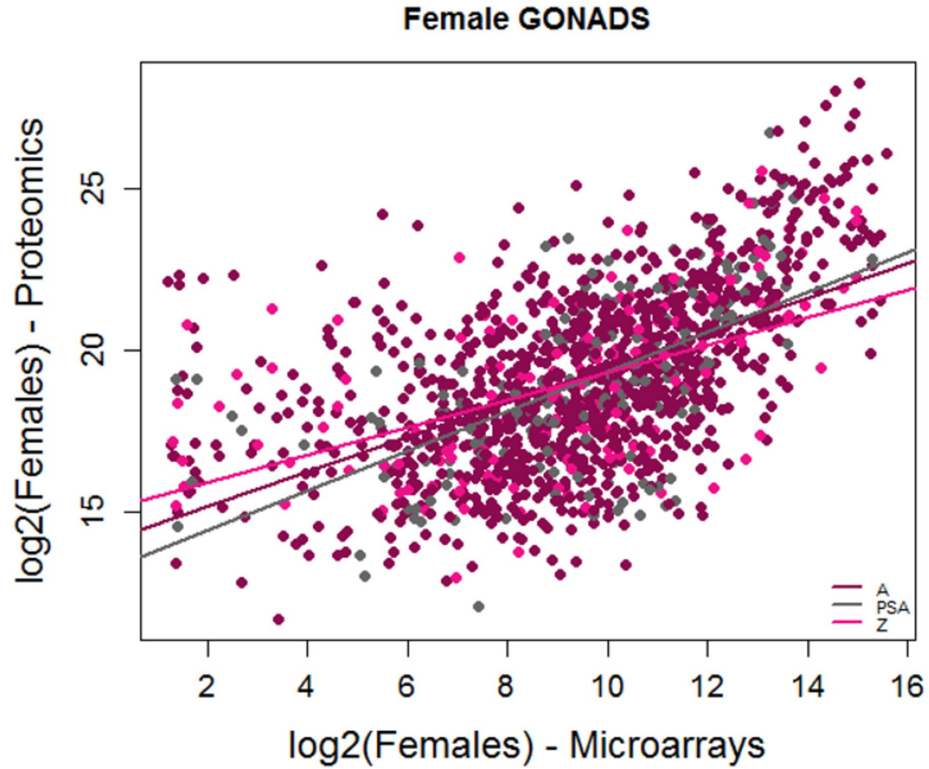
GONADS
A = 1041 genes
PSA = 159 genes
Z = 163 genes

R(A) = 0.54***
R(PSA) = 0.56 ***
R(Z) = 0.47***



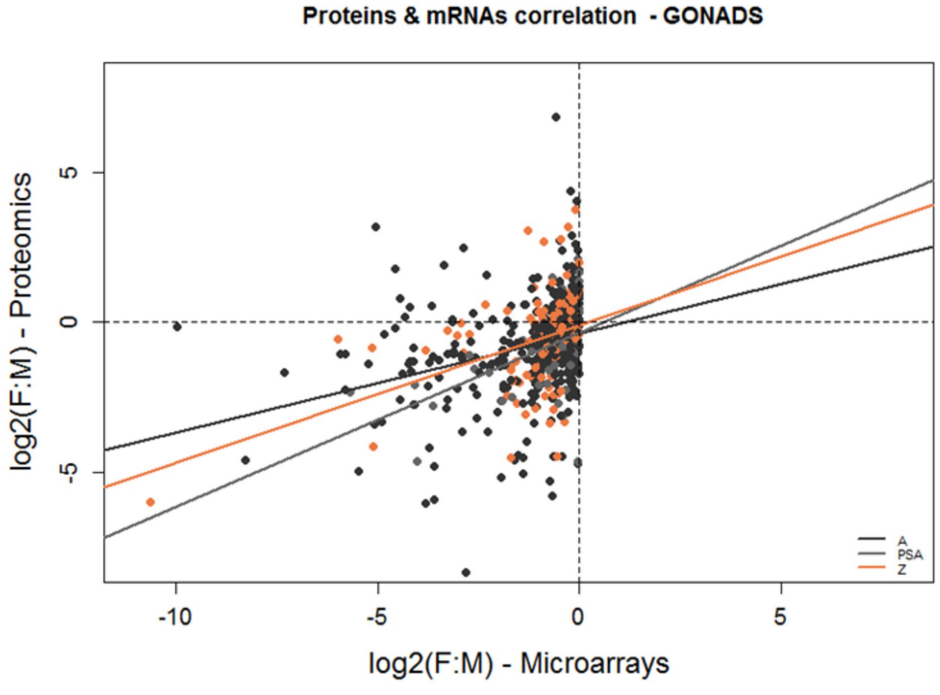
HEADS
A = 738 genes
PSA = 108 genes
Z = 123 genes

R(A) = 0.52***
R(PSA) =0.39***
R(Z) = 0.43***



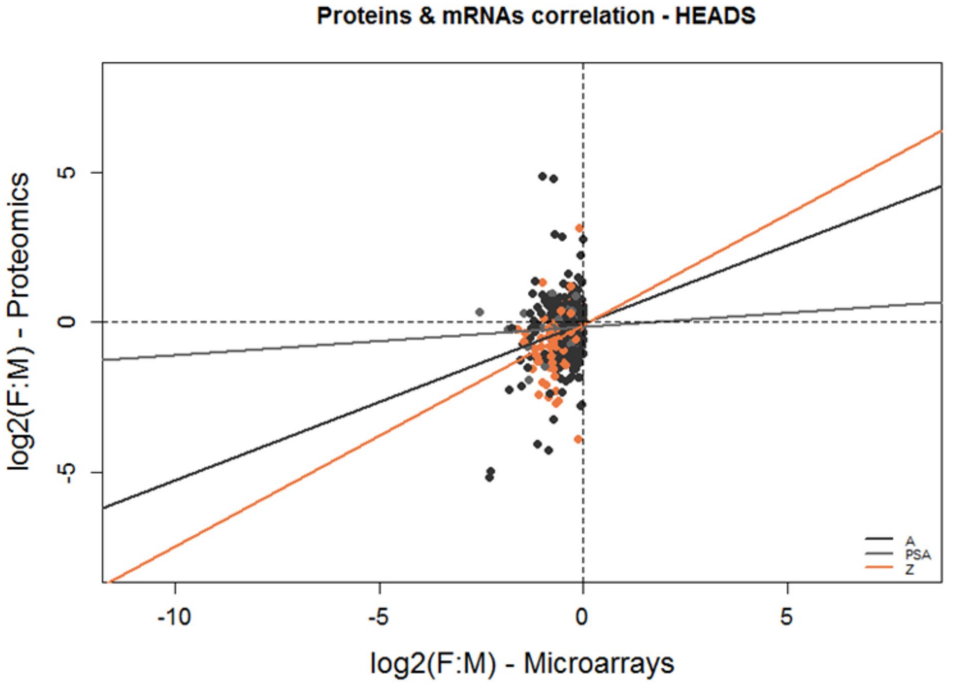
GONADS
A = 1041 genes
PSA = 159 genes
Z = 163 genes

R(A) = 0.55***
R(PSA) = 0.60***
R(Z) =0.53***



HEADS
A = 410 genes
PSA = 52 genes
Z = 125 genes

R(A) = 0.30***
R(PSA) = 0.48**
R(Z) =0.41***



GONADS
A = 339 genes
PSA = 49 genes
Z = 94 genes

R(A) = 0.20**
R(PSA) =0.08 N.S.
R(Z) =0.25*