Genes and pathways driving the maturation of *S. mansoni* after mating

UE Applied data analysis - supervised by Dr. Bob Zimmermann

Lukas Weilguny (01345185)

2 August 2018

Contents

1 Introduction

Schistosomes are trematodes of the genus *Schistosoma* that mature as parasites in veins of mammals and cause severe inflammation (King 2010). This is called Schistosomiasis and is the second most severe parasitic disease after malaria (Lu et al. 2016), transmitted as water-borne infectious disease after its larval stage in an intermediate snail host. Since eggs produced by adult worms after mating are the causative agent, understanding egg formation and maturation by mating, enabling the production of the pathogenic eggs, is very important. After migration of the immature worms to the portal vein of the final vertebrate host, they develop into adults with separate sexes. Thus, dioecy is a defining characteristic of *Schistosoma* within the phylum *Platyhelminthes*, which is otherwise composed of hermaphrodites (Lu et al. 2016). The mating pairs, which stay permanently physically coupled, then migrate to their final location, the mesenteric veins of the gut and subsequently egg production is initiated. Around one half of the eggs are shed into the gut lumen, whereas the other half are carried into the liver or spleen via the blood system, where they penetrate tissues to cause severe inflammation and liver cirrhosis (Lu et al. 2016).

Only one drug exists to treat this disease and resistance might be emerging, therefore new targets for drug development are needed. First insights into development and differentiation of the female gonad have been obtained after sequencing the genome of *S. mansoni*, nevertheless the processes involved are still poorly understood. Lu et al. (2016) thus performed tissue-specific RNA-seq analysis to identify molecules and mechanisms driving sexual maturation. To achieve that, they sequenced ovaries and testes of *S. mansoni* females and males from single-sex infections, resulting in undeveloped worms or bisex (mixed sex) infections, ensuing full sexual development of adult worms.

The work of Lu et al. (2016) is largely focused on identifying processes in the gonads and uses whole-worm samples as control. Instead, this project aims to reanalyze the data (Lu et al. 2017), specifically samples of whole worms, in order to detect the difference in transcriptome profiles of immature and mature adult worms and identify specific differentially expressed genes that drive the development of separate sexes. Additionally, by performing enrichment analysis we highlight cellular and metabolic processes that are active or disrupted upon maturation.

2 Material and Methods

2.1 Data source and pre-processing

The RNA-seq data used in this study is described in Lu et al. (2017), and deposited in the European Nucleotide Archive with the study accession number PRJEB14695. Full records on schistosome maintenance, worm isolation, nucleic acid extraction and RNA sequencing can be found in Lu et al. (2016). 100 ng of total RNA was used for library preparation in each sample with TruSeq RNA Library Preparation Kit. Sequencing was then performed in a multiplexed arrangement with unique Tags, resulting in 100 bp paired-end reads on an Illumina HiSeq 2500. The amount of reads corresponds to about 100x coverage of the S. mansoni transcriptome, which has an approximate size of 16 Mb.

The subset of sequenced specimens used for our investigation consists of 12 samples from whole adult worms, where 6 samples (3 male and 3 female) were obtained after single-sex infection of the final host and another 6 samples were collected from mixed-sex infected final hosts. The data was downloaded with the enaDataGet program from enaBrowserTools (v1.5.2) specifying the fastq format and default parameters otherwise. Further, individual files corresponding to one multiplexed sample were merged with standard Unix commands.

As a deviation from standard RNA-seq analysis reads were not trimmed, however a local alignment tool was chosen such that technical sequences would get soft clipped during mapping. The genome of *S. mansoni* and the associated transcriptome were obtained from EnsemblMetazoa in its assembly version ASM23792v2 (Protasio et al. 2012). The genome has an approximate size of 270 Mb with seven autosomes and two sex chromosomes (females: ZW, males: ZZ). The transcriptome was converted from GFF3 format to GTF with `gffread` included in cufflinks (Trapnell et al. (2010), v2.2.1) using default settings.

2.2 Read mapping and assignment

For mapping reads to the genome the STAR RNA-seq aligner (Dobin et al. (2013), v2.5.3a) was employed. Thus, a genome index was created prior to mapping with default parameters apart from `genomeSAindexNbases`, which was set to 6 in order to

reduces the length of the suffix array pre-indexing string for small genomes. Mapping was subsequently performed with predefined settings and standard parameters used in RNA-seq pipelines of the ENCODE project, which are listed in the manual of STAR (Dobin 2018). Afterwards, reads were assigned to genes with featureCounts (Liao et al. (2014), v1.6.2). Settings were changed to count fragments defined by paired-mates (`-p`) and to incorporate fractionally weighted multi-mapped reads (`-M --fraction`).

2.3 Differential gene expression analysis

SARtools (Varet (2016), v1.6.3) was applied for differential gene expression analysis, specifically utilizing DESeq2 (Love et al. 2014). The function `loadCountData` had to be modified to round fractional counts from multi-mapped reads. Additionally, the significance threshold was set to $alpha = 0.01$, after Benjamini-Hochberg adjustment.

2.4 Functional annotation

Functional annotation was performed using emapper-0.99 (Huerta-Cepas et al. 2017), with protein alignments produced by DIAMOND (Buchfink et al. 2014). Sequence searches were based on eggNOG orthologous groups (Huerta-Cepas et al. 2016), with default parameters, automatically adjusted taxonomic scope and targeting all orthologs in their database. DIAMOND was run in blastp-mode with the `--more_sensitive` flag, and stringent settings including an e-value cut-off of 0.001 and `--top` set to 3, meaning that only alignments within 3% of the top alignment score for each query are reported.

2.5 GO term enrichment analysis

Following differential gene expression analysis, a list of genes was selected according to the criterion of membership in the union of DEGs from the comparisons ms versus mm and fs versus fm, but not in the intersection of this set with DEGs from the comparison of ms versus fs. For a graphical representation please see the results section. This way, genes that were differentially regulated during male or female maturation, independent of general gender specific effects, were selected. This set was subsequently analyzed for enriched GO terms found by functional annotation using topGO (Alexa

Table 1: The number of raw reads, uniquely and multi-mapped reads as well as the total sequences assigned to genes

Sample	raw reads	uniquely mapped (n, %)		multi mapped (n, %)		assigned reads (n, %)	
fs1	31537628	26545508	84.17	1504922	4.77	24805324	78.65
fs2	55985134	45678263	81.59	2988914	5.34	43027141	76.85
fs3	40615071	29850099	73.5	1598788	3.94	27839307	68.54
fm1	41259411	33950282	82.28	2497677	6.05	30792553	74.63
fm2	28525293	18180409	63.73	5781728	20.27	21276807	74.59
fm3	24373230	15589983	63.96	5228620	21.45	18342923	75.26
mm1	21126517	10405541	49.25	8386109	39.69	11971967	56.67
mm2	30864815	24866495	80.57	1409732	4.57	22925174	74.28
mm3	29560640	24514810	82.93	1506957	5.1	22818931	77.19
ms1	31844449	26128211	82.05	1543312	4.85	24370884	76.53
ms2	32127433	26027564	81.01	1514551	4.71	24267282	75.53
ms3	27276243	22775186	83.5	1395874	5.12	20980196	76.92
Mean	32924655.33	25376029.25	75.71	2946432	10.49	24451540.75	73.8
S.d.	9260198.36	8982298.96	11.04	2292109.92	11.07	7489891.92	5.95

and Rahnenfuhrer 2016), with the classicCount algorithm and Fisher's Exact test as well as the GO hierarchy of the GO.db package. Visualizations were produced with GOplot (Walter et al. 2015).

3 Results

3.1 Read statistics and diagnostics

The present study examines differentially expressed features between samples of two conditions and two sexes of *S. mansoni*. An overview of the samples, their raw read counts, mapping statistics and counted features is given in Table 1. The percentage of null counts for each sample are shown in Fig. 1, which are consistent between conditions and also within replicates. Notably 2808 features or 21.09 % have no counts across all samples, partly explicable by the transcriptome containing nuclear as well as mitochondrial rRNAs and tRNAs. Fig. 1B further shows the density of read counts, again the concordance between samples is high. However, the overall shape does not follow an expected Poisson distribution but instead shows a bias of features with rather high counts over genes with medium or low counts. This pattern was also observed by the authors of the published data (see Fig. 3, Lu et al. 2017).

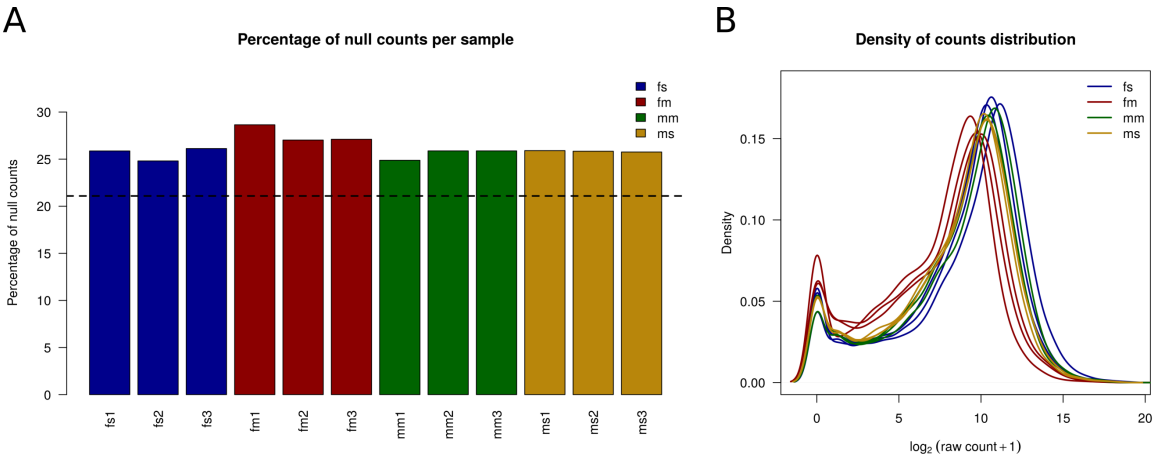


Figure 1: (A) The percentage of null counts is consistent across samples and replicates, yet a rather high percentage of features in the gene annotation has no counts in any of the samples (dashed line). (B) The density of read counts also shows consistency across all samples but does not follow an expected Poisson distribution. Instead a bias for high-count features over medium and low counts can be observed.

3.2 Assessing variability within and across samples

Fig. 2A illustrates the percentage of counts that are assigned to the most abundant feature. Apart from the fm samples, all other conditions show values below 4% and thus absence of high count features. Strikingly, for all three female replicates of mixed-infection experiments a hypothetical, uncharacterized gene makes up 10-25% of all assigned reads, indicative of a low-complexity sample or some technical issue.

To further assess the similarity of replicates and whether the captured variation has biological reasons, the first three principal components are visualized (Fig. 2B and 2C). Generally the replicates within their conditions are clustered together. What is more, the first principal component separates the samples by biological condition with 81% of the total variance explained. This axis also shows that male *S. mansoni* prior to maturation (ms) do not seem to differ from their matured counterparts (mm) as much as females do after egg production has been initiated. Additionally, females before maturation (fs) appear to be more similar to male worms than they are to matured female *S. mansoni*. This is also clearly visible in Fig. 2D, where samples were hierarchically clustered using the euclidean distance after VST-transformation. Female worms of single-sex infection experiments cluster with both conditions of male samples in support of the results from

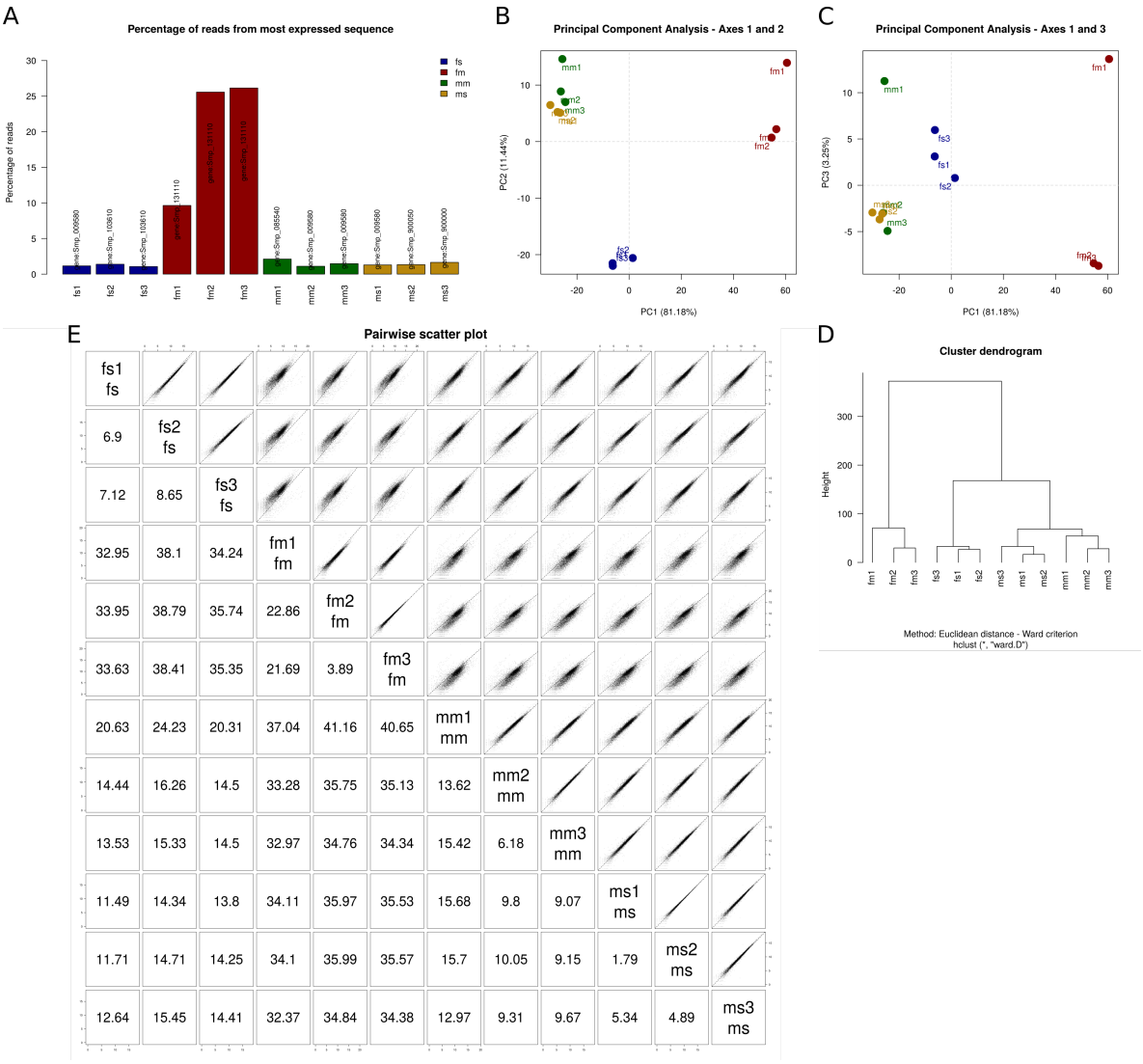


Figure 2: Variability of investigated samples. (A) For all replicates of condition *fm* 10-25% of mapped reads were assigned to a single feature; a hypothetical, uncharacterized gene. Whereas the most abundant feature of the remaining samples makes up less than 4% of reads. (B, C) Principal component analysis reveals good clustering of replicates and separates samples by their biological condition. (D) Hierarchical clustering additionally groups replicates into subclades and together with PCA hints at a higher similarity of immature female *S. mansoni* to their male counterparts. (E) A pairwise scatter plot shows feature counts between all pairwise combinations of samples. It confirms the quality of replicates for most conditions but also highlights the dissimilarity of some samples to their biological replicates (e.g. fm1).

PCA. Besides that, replicates of conditions group in subclades as expected for biological replicates.

Similar patterns of resemblance can be observed in a pairwise scatter plot (Fig. 2E). The upper triangle shows scatter plots of $\log_2$-transformed counts between samples and the lower triangle lists the SERE statistic, which behaves inversely proportional to the similarity between samples. Notably, most replicates exhibit good resemblance with the closest pair separated by a SERE value of 1.79 (ms1 and ms2). At the same time however, sample fs1 seems fairly dissimilar to its biological replicates (e.g. 22.86 to fs2). This could partly be explained by the apparent low complexity of the fm samples, as seen in Fig. 2A, and possibly the influence of that major feature on the dissimilarity. A general trend, which further emphasizes the patterns seen by PCA and clustering is that mature females are the least similar to all other samples, whereas immature females are more similar to males irrespective of their developmental stage (see upper right corner and lower left corner of Fig. 2E).

3.3 Normalization and dispersion estimation

To make read counts comparable across samples, normalization was performed with the DESeq2 method (Love et al. 2014). Boxplots in Fig. 3A and 3B visualize the distribution of reads per sample prior to and after normalization, respectively. The normalized distributions in Fig. 3B are clearly stabilized across all conditions and replicates. Another step in modelling differential gene expression in DEseq2 involves the estimation of dispersion of the data. The relationship of the mean normalized counts and the associated variance fitted as a model to the empirical data can be seen in Fig. 3C. Additionally, a graphical check for the assumption of log-normality of the dispersion values is shown in Fig. 3D, which shows a slight shift of the main peak towards lower negative values and a second, albeit smaller peak at positive values.

3.4 Differentially expressed genes

The absolute number of differentially expressed genes (DEGs) is presented in Table 2. As expected from previous descriptive and diagnostic analyses, the comparisons involving mature female *S. mansoni* exhibit most DEGs (5235-5835) again confirming that these

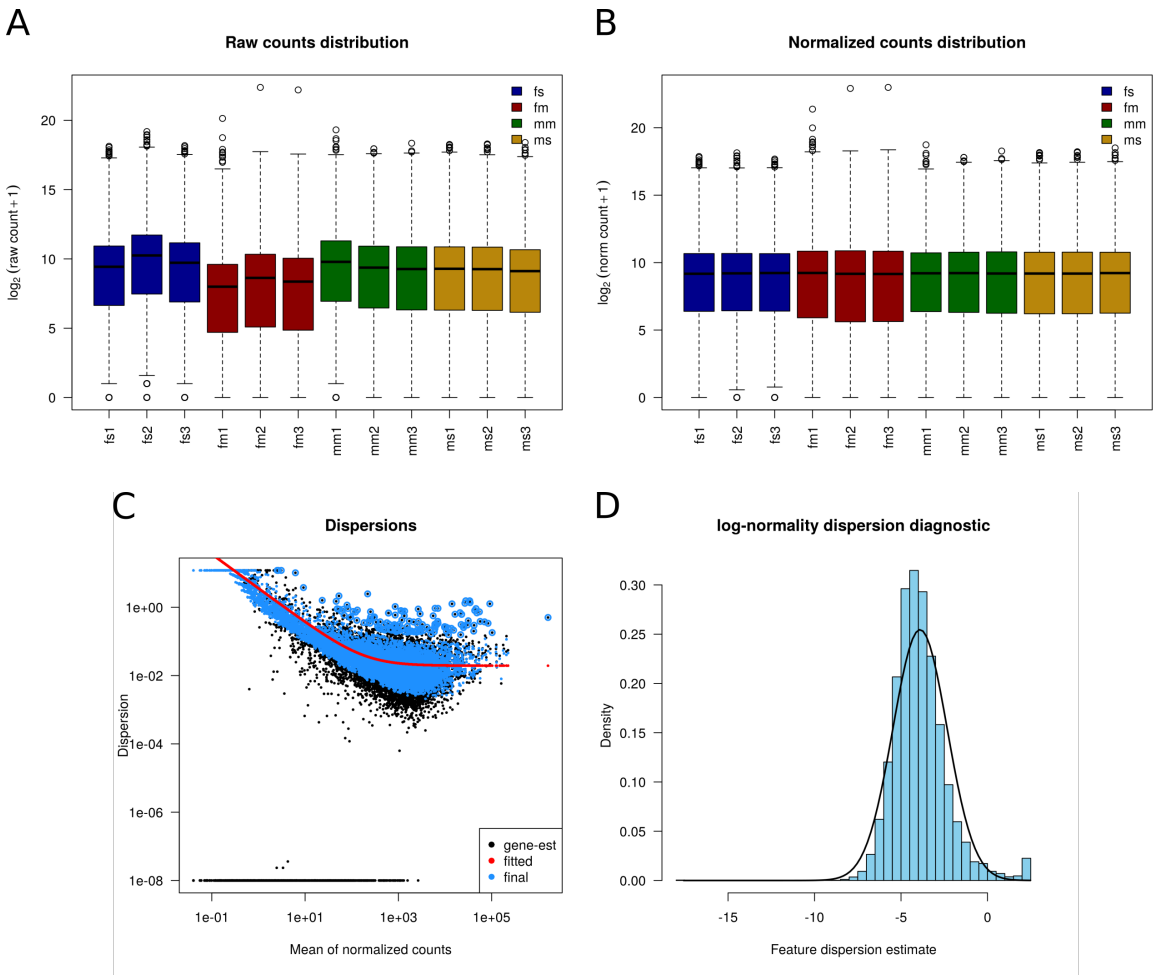


Figure 3: (A, B) Normalization with DESeq2 stabilizes the distribution of raw counts, in order to perform differential gene expression analysis across conditions. (C) The dispersion estimates as a relationship of mean normalized counts and variance as used in the DESeq2 modelling process. (D) The assumption of log-normality of estimated dispersion values is not fulfilled unconditionally.

Table 2: Summary of the number of differentially expressed genes between all tested conditions

Conditions	Down-regulated	Up-regulated	Total DEGs
fm vs fs	2616	2619	5235
mm vs fs	1613	1721	3334
ms vs fs	1247	1412	2659
mm vs fm	2820	3015	5835
ms vs fm	2755	2941	5696
ms vs mm	369	292	661

samples are the most dissimilar to all remaining ones. Whereas the undeveloped females fall in between with around 3000 DEGs, the least differential features were detected for immature versus mature males (661 in total). A graphical representation of the DEGs is given in Fig. 4A, which sets the numbers in relation to the overall number of features. Further, Fig. 4B shows the number of DEGs in each comparison on the bottom left and the size of all possible intersections on the top. It demonstrates that the largest intersection is obtained from the three largest sets, combining gender specific differences after maturation of the female and the process of female development itself. Moreover, we can see that the most private DEGs (276) are observed for the comparison of mm versus fm, both genders after maturation. These might include genes that are responsible for sex-specific functions after maturation, such as egg and sperm production, as opposed to the general, sex-independent process of maturation. Genes that are significantly differentially expressed across all conditions are rare with only 73 detected features.

In contrast to the overview provided in the first two panels, Fig. 4C shows the results of each comparison in greater detail. These plots present the log2-fold change as a function of normalized counts for each investigated feature, colored red in case of significance. Taken together, these results confirm that maturation has a much smaller impact on the transcriptional state of male worms than it has on females. Actually, the development of the female is characterized by a shift in gene regulation that is almost of the same intensity as the difference between the genders, and notably it encompasses many of the same genes as visible in the largest intersection of Fig. 4B.

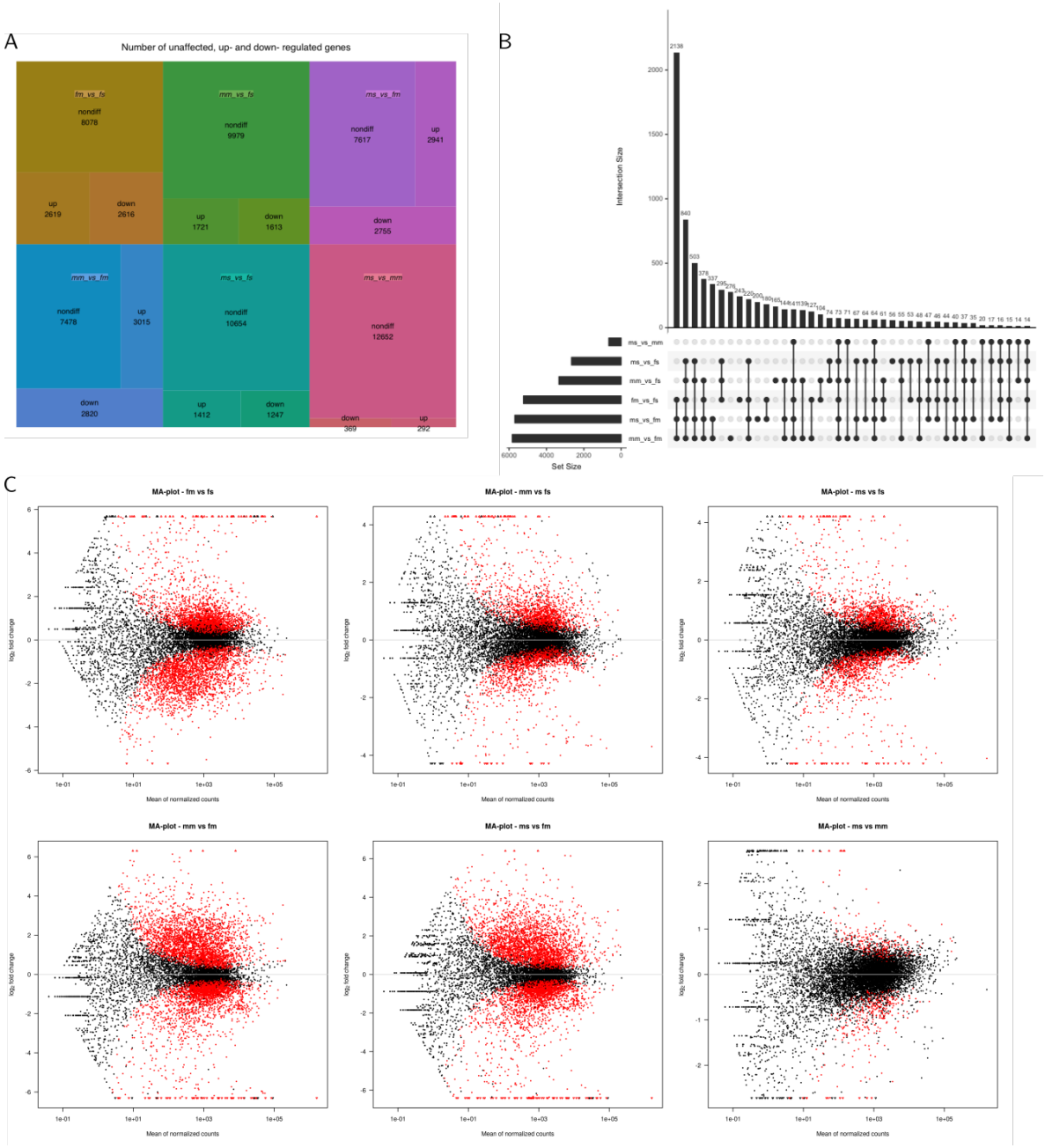


Figure 4: Differential gene expression analysis. (A) A treemap shows the proportion of up- and down- regulated as well as non-differentially expressed genes for each comparison. (B) The size of the intersections between comparisons of all conditions. (C) MA-plots show log2-fold changes as a function of normalized counts. Features are represented as dots and colored red to illustrate significance.

3.5 Selecting a subset of relevant genes

In order to further the understanding of the transition between hermaphroditic and dioecious reproductive states, Chong et al. (2013) investigated how simultaneous hermaphrodites, specify and maintain gender. They identified a male-specific gene (*dmd-1*) coding for Doublesex-and mab3-related transcription factor in *Schmidtea mediterranea*, and a homolog in *S. mansoni*, which acts as a major driver of sex transition and male germ cell development as one means by which modulation of sex-specific pathways can cause transition from hermaphrodites to dioecy. These new findings together with the possibility of evolution of dioecy through androdioecy, as proposed by Platt and Brooks (1997), point at an interesting subset of genes for further investigation. More specifically, since our data suggest an evolutionary history of protandry as a form of sequential hermaphroditism in the ancestor by showing that female virgins are more similar to males (Fig. 2) and that the transition of maturity is much more pronounced for females (Fig. 4) we propose in depth analysis of the subset of genes depicted in Fig. 5. These genes follow the same pattern as the homolog to *Smed_dmd-1*, reported in Chong et al. (2013), as they are differentially regulated in males or in females but not in a comparison of undeveloped females versus males, in order to exclude other general gender differences not directly involved in the developmental process. This list of genes was used in a GO term enrichment analysis using a custom functional annotation of the *S. mansoni* proteome.

3.6 GO terms enriched in the developmental process of *S. mansoni*

The bubble plot in Fig. 6 helps to get an overview of the enriched GO terms. A z-score is shown on the x-axis, which hints at either a down- or up- regulation of a GO term by considering the relative difference of down- and up- regulated genes associated with that GO term (see Walter et al. (2015) for details). Whereas on the y-axis, GO terms are plotted according to their negative logarithmic p-values. Generally we observe a shift of z-scores to positive values, indicating that more pathways or functions are being up- rather than down-regulated during the maturation process. Notably, significantly enriched GO terms involve many categories associated with development and growth as well as increased protein and membrane production and glycosylation (Fig. 6, right),

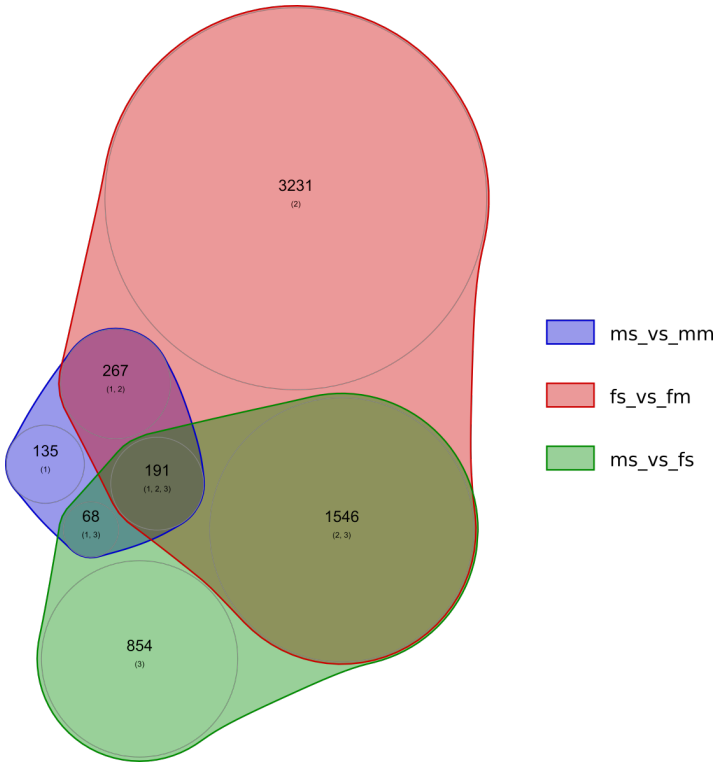
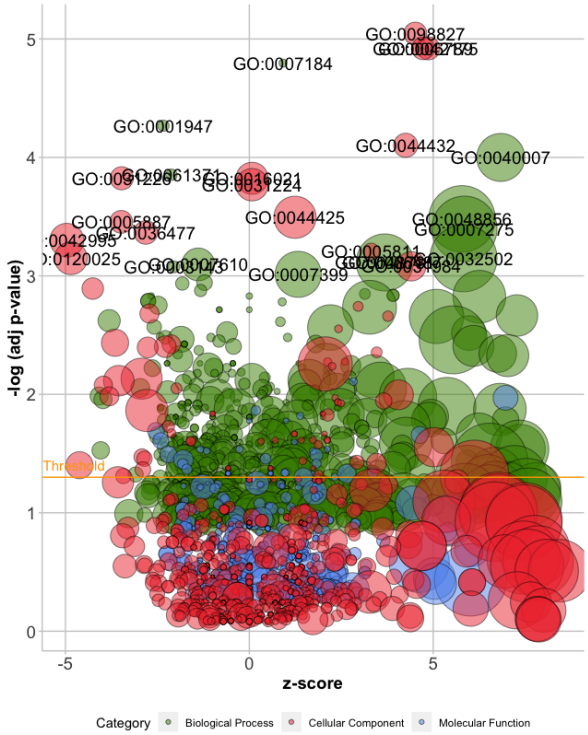


Figure 5: The genes selected for further analysis involve DEGs from the comparison of immature and mature worms of both sexes (union of red and blue) without DEGs from the comparison of immature males and females (intersection with green)

corresponding very well to the maturation process that is initiated after mating. An additional visualization is provided in Fig. 7, which shows the top 5 terms of Fig. 6 within the network of the GO hierarchy. This illustrates how the significant GO terms are distributed over the GO graph includes all subgraphs induced by these terms with lowest p-value.



ID	Description
GO:0007184	SMAD protein import into nucleus
GO:0001947	heart looping
GO:0040007	growth
GO:0061371	determination of heart left/right asymme...
GO:0048856	anatomical structure development
GO:0007275	multicellular organism development
GO:0032502	developmental process
GO:0006487	protein N-linked glycosylation
GO:0048731	system development
GO:0007610	behavior
GO:0003143	embryonic heart tube morphogenesis
GO:0007399	nervous system development
GO:0098827	endoplasmic reticulum subcompartment
GO:0042175	nuclear outer membrane-endoplasmic retic...
GO:0005789	endoplasmic reticulum membrane
GO:0044432	endoplasmic reticulum part
GO:0016021	integral component of membrane
GO:0031226	intrinsic component of plasma membrane
GO:0031224	intrinsic component of membrane
GO:0044425	membrane part
GO:0005887	integral component of plasma membrane
GO:0036477	somatodendritic compartment
GO:0042995	cell projection
GO:0005811	lipid droplet
GO:0120025	plasma membrane bounded cell projection
GO:0031984	organelle subcompartment

Figure 6: An overview of the GO enrichment analysis is gained by a bubble plot (Walter et al. 2015), where terms are arranged by their z-score and p-value. The area of each circle is proportional to the number of genes assigned to the respective GO term and colors indicate the category of the term. A significance threshold is drawn for alpha = 0.05. Circles with a negative logarithmic p-value greater than 3 are labelled by their GO-Ids and shown in the table to the right.

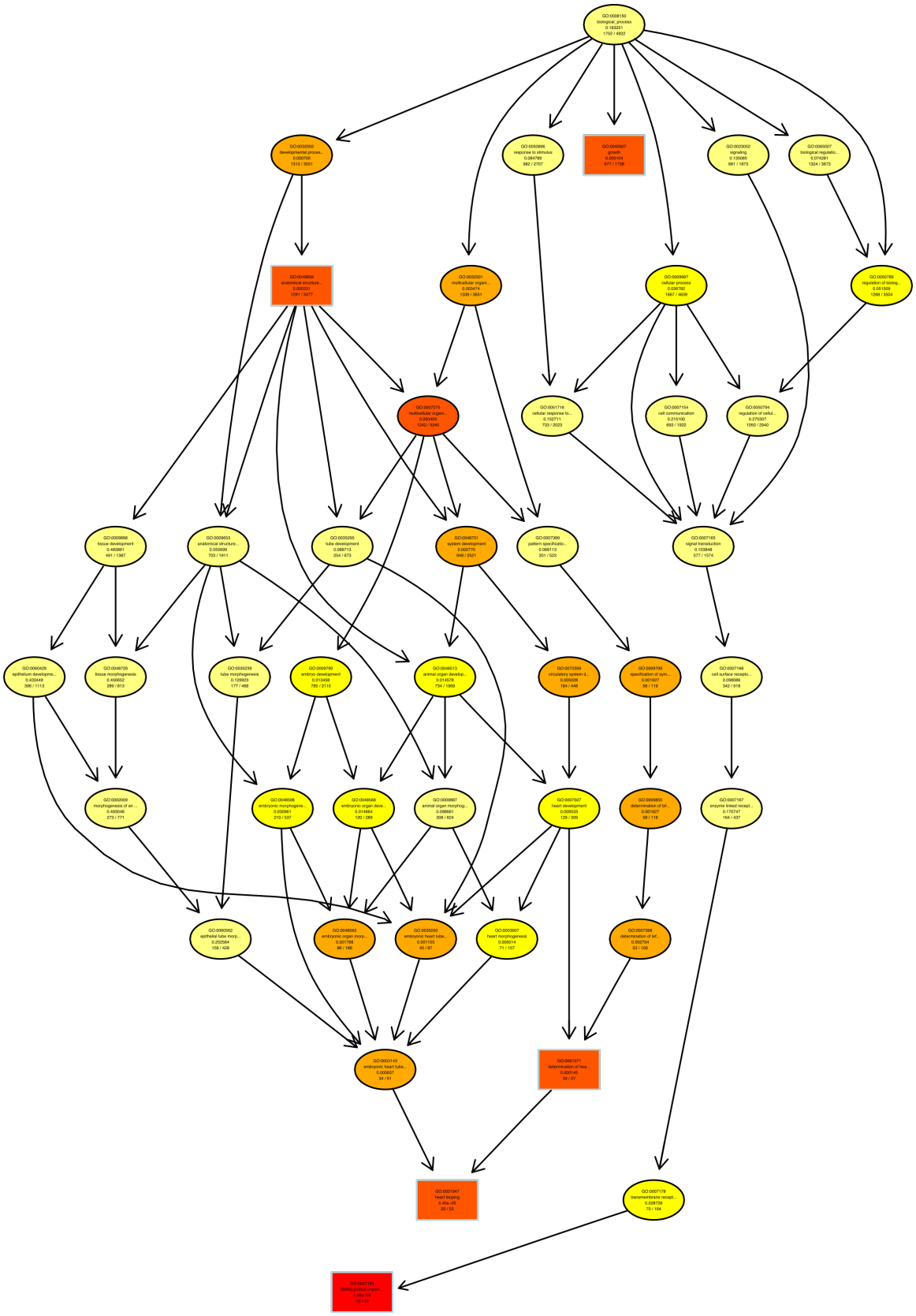


Figure 7: The five terms with lowest p-value in Fig. 6 are depicted within their GO hierarchy as rectangles. Each GO term is colored on a gradient according to its p-value.

4 Discussion and conclusion

In this work we reanalyzed a dataset published by (Lu et al. 2016) on the transcriptomic profiles of immature and fully developed, adult *S. mansoni*. RNA-seq data was processed and analyzed according to standard pipelines to assess variability and similarity between different conditions and specifically to detect differentially expressed genes. We successfully show that the samples capture true biological variation and replicates share a reasonable degree of similarity and complexity. One possible case of low complexity was observed in female samples of mixed-infection experiments. In particular, one hypothetical, uncharacterized gene makes up 10-25% of all assigned reads, indicative of low-complexity or a technical issue. However, fitting for the condition of mixed-infection, a Blast search with the translated protein sequence produced a hit to the eggshell protein gene region of *S. haematobium*, a sister species to *S. mansoni*. Thus, the extraordinary high expression of that gene could be related to the massive daily production of up to 3000 eggs per mature female (Lu et al. 2016).

Another notable result is the similarity of undeveloped females to male worms, which was already noticed by Lu et al. (2016). We can confirm their findings by PCA, hierarchical clustering and by pair-wise scatterplots, demonstrating that the transcriptional profile of immature females resembles the profile of male worms more closely than of adult females. A possible explanation for this striking outcome is the evolutionary context of the schistosome gender. In hermaphroditic worms of the phylum *Platyhelminthes*, including cestodes and trematodes amongst others, protandry is a common form of sequential development, in which the female gonads form after the male reproductive organs. By contrast, in schistosomes dioecy is thought to have evolved for partition of labor in reproduction and egg production (Platt and Brooks 1997). This occurs for example by genes suppressing female-specific functions in males and thus prohibiting further maturation of female gonads (Chong et al. 2013). Together with the permanent contact after pairing, reminiscent of a protandric ancestor, this hints at *S. mansoni* inhabiting a transitional state towards fully established dioecy. Additional evidence stems from the detection of leaky transcription of genes specifically for egg production in male worms (Chong et al. 2013).

With regard to potential evolutionary novelties for the origin of dioecy, we investigated a set of differentially expressed genes between undeveloped and adult stages, while excluding gender specific functions decoupled of the developmental process. The sig-

nificantly enriched GO terms of this subset of genes includes developmental processes, anatomical structure development, integral components of the membrane and additional categories generally associated with growth. Taken together, we propose this subset of candidate genes for further investigation of the maturation process and its evolutionary history in the context of the origin of dioecy. Follow-up studies might assess orthology relationships of these genes in a wider context, covering the different mating or gender systems at the base of bilaterian animals. Moreover, knock-out experiments targeting developmental pathways involving such genes might help to identify potential targets for prevention of maturation during schistosome infection.

5 References

Alexa A, and Rahnenfuhrer J. 2016. TopGO: Enrichment analysis for gene ontology. *R bioconductor package version 2.32.0.*

Buchfink B, Xie C, and Huson D. 2014. Fast and sensitive protein alignment using diamond. *Nature Methods.* 12(1): 59–60.

Chong T, Collins J, Brubacher J, Zarkower D, and Newmark P. 2013. A sex-specific transcription factor controls male identity in a simultaneous hermaphrodite. *Nature Communications.* 4(1814):.

Dobin A. 2018. STAR manual 2.6.0a.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR. 2013. STAR: Ultrafast universal rna-seq aligner. *Bioinformatics.* 29(1): 15–21.

Huerta-Cepas J, Forslund K, Coelho L, Szklarczyk D, Jensen L, Von Mering C, and Bork P. 2017. Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Molecular Biology and Evolution.* 34(8): 2115–2122.

Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter M, Rattei T, Mende D, Sunagawa S, Kuhn M, Jensen L, Von Mering C, and Bork P. 2016. EGGNOG 4.5: A hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Research.* 44(D1): D286–D293.

King C. 2010. Parasites and poverty: The case of schistosomiasis. *Acta Tropica.* 113(2): 95–104.

Liao Y, Smyth GK, and Shi W. 2014. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* 30(7): 923–930.

Love MI, Huber W, and Anders S. 2014. Moderated estimation of fold change and dispersion for rna-seq data with deseq. *Genome Biology.* 15(12): 550.

Lu Z, Sessler F, Holroyd N, Hahnel S, Quack T, Berriman M, and Grevelding C. 2017. Data descriptor: A gene expression atlas of adult schistosoma mansoni and their

gonads. *Scientific Data*. 4(170118):.

Lu Z, Sessler F, Holroyd N, Hahnel S, Quack T, Berriman M, and Grevelding C. 2016. Schistosome sex matters: A deep view into gonad-specific and pairing-dependent transcriptomes reveals a complex gender interplay. *Scientific Reports*. 6(31150):.

Platt T, and Brooks D. 1997. Evolution of the schistosomes (digenea: Schistosomatoidea): The origin of dioecy and colonization of the venous system. *Journal of Parasitology*. 83(6): 1035–1044.

Protasio AV, Tsai IJ, Babbage A, Nichol S, Hunt M, Aslett MA, De Silva N, Velarde GS, Anderson TJC, Clark RC, Davidson C, Dillon GP, Holroyd NE, LoVerde PT, Lloyd C, McQuillan J, Oliveira G, Otto TD, Parker-Manuel SJ, Quail MA, Wilson RA, Zerlotini A, Dunne DW, and Berriman M. 2012. A systematically improved high quality genome and transcriptome of the human blood fluke schistosoma mansoni. *PLOS Neglected Tropical Diseases*. 6(1): 1–13.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, Baren MJ van, Salzberg SL, Wold BJ, and Pachter L. 2010. Transcript assembly and quantification by rna-seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*. 28511–515.

Varet LAC Hugo AND Brillet-Guéguen. 2016. SARTools: A deseq2- and edger-based r pipeline for comprehensive differential analysis of rna-seq data. *PLOS ONE*. 11(6): 1–8.

Walter W, Sánchez-Cabo F, and Ricote M. 2015. GOplot: An r package for visually combining expression data with functional analysis. *Bioinformatics*. 31(17): 2912–2914.