

METHODOLOGY ARTICLE

Open Access

Efficient depletion of ribosomal RNA for RNA sequencing in planarians



Iana V. Kim^{1*}, Eric J. Ross^{2,3}, Sascha Dietrich⁴, Kristina Döring⁴, Alejandro Sánchez Alvarado^{2,3} and Claus-D. Kuhn^{1*} 

Abstract

Background: The astounding regenerative abilities of planarian flatworms prompt steadily growing interest in examining their molecular foundation. Planarian regeneration was found to require hundreds of genes and is hence a complex process. Thus, RNA interference followed by transcriptome-wide gene expression analysis by RNA-seq is a popular technique to study the impact of any particular planarian gene on regeneration. Typically, the removal of ribosomal RNA (rRNA) is the first step of all RNA-seq library preparation protocols. To date, rRNA removal in planarians was primarily achieved by the enrichment of polyadenylated (poly(A)) transcripts. However, to better reflect transcriptome dynamics and to cover also non-poly(A) transcripts, a procedure for the targeted removal of rRNA in planarians is needed.

Results: In this study, we describe a workflow for the efficient depletion of rRNA in the planarian model species *S. mediterranea*. Our protocol is based on subtractive hybridization using organism-specific probes. Importantly, the designed probes also deplete rRNA of other freshwater triclad families, a fact that considerably broadens the applicability of our protocol. We tested our approach on total RNA isolated from stem cells (termed neoblasts) of *S. mediterranea* and compared ribodepleted libraries with publicly available poly(A)-enriched ones. Overall, mRNA levels after ribodepletion were consistent with poly(A) libraries. However, ribodepleted libraries revealed higher transcript levels for transposable elements and histone mRNAs that remained underrepresented in poly(A) libraries. As neoblasts experience high transposon activity this suggests that ribodepleted libraries better reflect the transcriptional dynamics of planarian stem cells. Furthermore, the presented ribodepletion procedure was successfully expanded to the removal of ribosomal RNA from the gram-negative bacterium *Salmonella typhimurium*.

Conclusions: The ribodepletion protocol presented here ensures the efficient rRNA removal from low input total planarian RNA, which can be further processed for RNA-seq applications. Resulting libraries contain less than 2% rRNA. Moreover, for a cost-effective and efficient removal of rRNA prior to sequencing applications our procedure might be adapted to any prokaryotic or eukaryotic species of choice.

Keywords: Planarians, *Schmidtea mediterranea*, Ribosomal RNA removal, rRNA depletion, RNA sequencing

Background

Freshwater planarians of the species *Schmidtea mediterranea* are well known for their extraordinary ability to regenerate. This ability is supported by the presence of a large population of adult pluripotent stem cells, termed neoblasts [1]. Neoblasts are capable of producing all planarian cell types [2]. Moreover, they preserve their potency over the whole lifespan of the animal, which seems to be infinite [3]. Therefore, planarians embody

an excellent model to study regeneration, aging and stem cell-based diseases. The phylum Platyhelminthes, to which *S. mediterranea* belongs, includes multiple other members that display varying degrees of regenerative abilities. While some freshwater species (e.g. *Dugesia japonica* and *Polycelis nigra*) are capable to restore their body from any tiny piece [4, 5], others (e.g. *Procotyla fluviatilis*) have limited anterior regeneration abilities [6]. Altogether, the ability to regenerate seems not solely based on the presence of pluripotent stem cells, but represents a complex interplay between different signaling pathways. The underlying changes in gene expression

* Correspondence: iana.kim@uni-bayreuth.de; claus.kuhn@uni-bayreuth.de

¹Gene regulation by Non-coding RNA, Elite Network of Bavaria and University of Bayreuth, Universitätsstrasse 30, 95447 Bayreuth, Germany
Full list of author information is available at the end of the article



therefore need to be studied using transcriptome-wide techniques like RNA sequencing.

For any informative RNA-seq library preparation, ribosomal RNA, comprising > 80% of total RNA, has to be removed. To achieve this goal two strategies can be pursued: either polyadenylated (poly(A)) RNA transcripts are enriched or rRNA is removed. Both approaches have advantages and limitations. On the one hand, the enrichment of poly(A) transcripts ensures better coverage of coding genes compared to ribodepleted samples, when sequenced to similar depth [7]. However, this advantage is outweighed by the loss of transcripts lacking poly(A) tails, which include preprocessed RNAs, a large share of all non-coding RNAs, such as enhancer RNAs and other long non-coding RNAs. In addition, long terminal repeat (LTR) retrotransposons and various intermediates of endonucleolytic RNA degradation are lost during poly(A) selection [8–13]. Furthermore, most prokaryotic RNAs lack poly(A) tails, making rRNA depletion crucial for the study of bacterial transcriptomes [14].

Here, we describe a probe-based subtractive hybridization workflow for rRNA depletion that efficiently removes planarian rRNA from total RNA. The protocol can be applied to input as low as 100 ng total RNA, which corresponds to 100,000 FACS-sorted planarian stem cells (X1 population) [15, 16]. Moreover, the DNA probes developed for *S. mediterranea* were successfully used for the removal of ribosomal RNA in related planarian species of the order Tricladida. The rRNA removal workflow presented here is also easily adapted to other organisms, as demonstrated by the removal of rRNA from total RNA of *Salmonella typhimurium* using organism-specific probes.

Results

Development of an efficient rRNA depletion protocol for planarians

To deplete ribosomal RNA from planarian total RNA, we chose to develop a protocol based on the hybridization of rRNA-specific biotinylated DNA probes to ribosomal RNA and the capture of the resulting biotinylated rRNA-DNA hybrids by use of streptavidin-coated magnetic beads (Fig. 1a). To that end, we synthesized a pool of 88 3'-biotinylated 40-nt long DNA oligonucleotide probes (siTOOLS Biotech, Martinsried, Germany). We chose probes with a length of 40 nucleotides since their melting temperature in DNA-RNA hybrids was shown to be $80 \pm 6.4^\circ\text{C}$ in the presence of 500 mM sodium ions [17]. This would allow probe annealing at 68°C in agreement with generally used hybridization temperatures [18]. The probes were devised in antisense orientation to the following planarian rRNA species: 28S, 18S type I and type II, 16S, 12S, 5S, 5.8S, internal transcribed spacer (ITS) 1 and ITS 2 (Additional file 1).

To assess RNA quality and the efficiency of rRNA removal, we used capillary electrophoresis (Fragment Analyzer, Agilent). The separation profile of total planarian RNA only shows a single rRNA peak at about 1500 nucleotides (nts) (Fig. 1b). This single rRNA peak is the result of the 28S rRNA being processed into two fragments that co-migrate with the peak of 18S rRNA [19]. Planarian 28S rRNA processing usually entails the removal of a short sequence located in the D7a expansion segment of 28S rRNA. The length of the removed fragment thereby varies between 4 nts and 350 nts amongst species (e.g. in *Dugesia japonica* 42 nts are removed) [19]. Intriguingly, a similar rRNA maturation process was observed in particular protostomes, in insects such as *D. melanogaster* and in other Platyhelminthes [19–21]. In addition to the 28S rRNA maturation phenomenon, *S. mediterranea* possesses two 18S rDNA copies that differ in about 8% of their sequence. However, only 18S rRNA type I was reported to be functional and predominantly transcribed [22, 23].

As a first step during rRNA removal all 88 DNA probes were annealed to total planarian RNA. Since RNA molecules are negatively charged, the presence of cations facilitates the annealing of probes to RNA by reducing the repulsion of phosphate groups [24, 25]. Although Mg^{2+} ions are most effective in stabilizing the tertiary structure of RNA and in promoting the formation of DNA-RNA hybrids, they are also cofactors for multiple RNases [26] and hence should not be included during ribodepletion. Therefore, we tested several hybridization buffers with varying concentrations of sodium ions (Fig. 1c). In the absence of sodium ions we could only accomplish an incomplete removal of rRNA. However, hybridization buffers with a sodium concentration > 250 mM led to the complete depletion of rRNA from planarian total RNA (Fig. 1c, d). Thus, optimal rRNA removal requires the presence of > 250 mM NaCl in the hybridization buffer. As we obtained the most consistent results in the presence of 500 mM NaCl, we decided to utilize this salt concentration in our procedure (Fig. 1d).

Detailed rRNA depletion workflow

Required buffers

Hybridization buffer (20 mM Tris-HCl (pH 8.0), 1 M NaCl, 2 mM EDTA).

Solution A (100 mM NaOH, 50 mM NaCl, DEPC-treated).

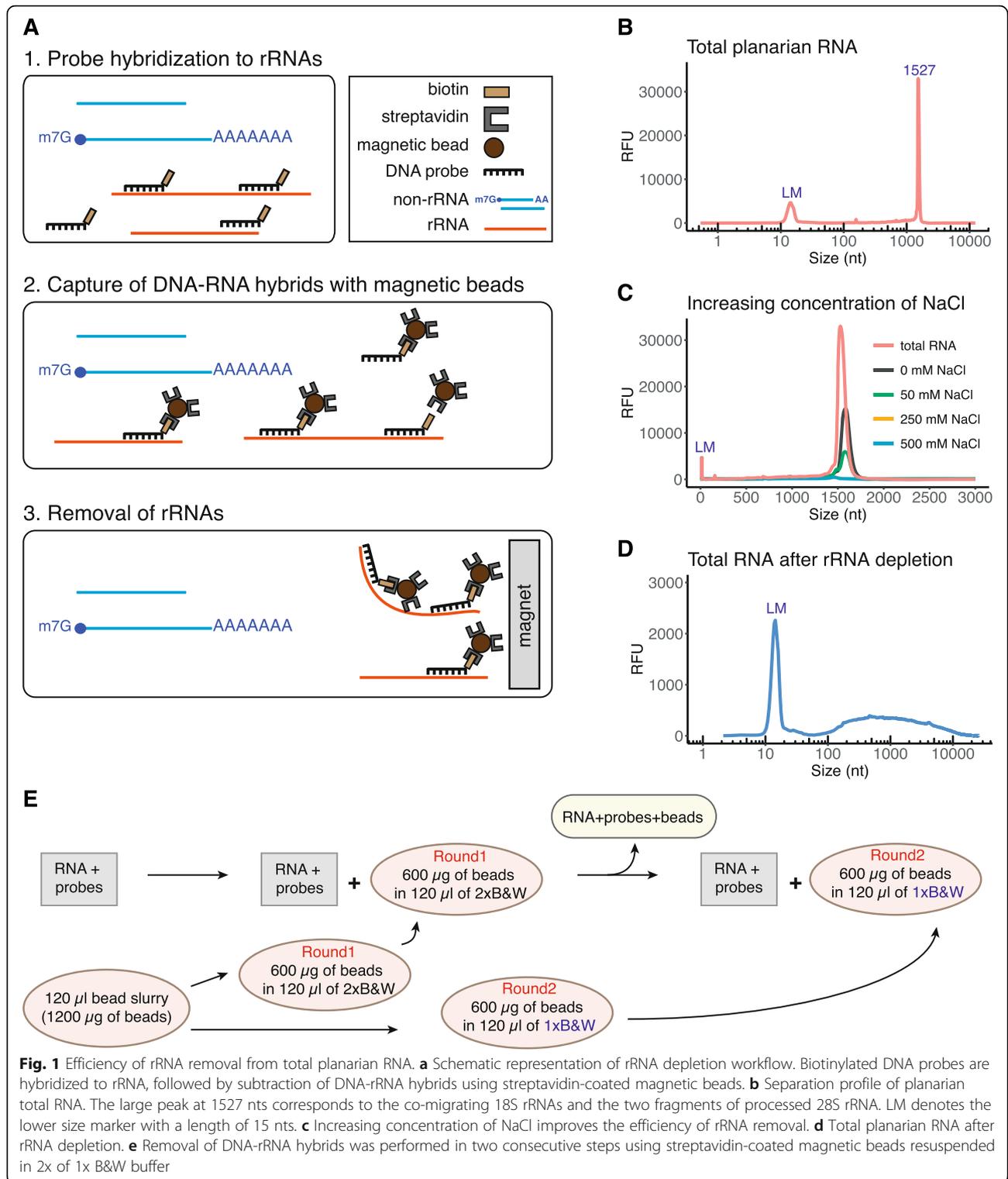
Solution B (100 mM NaCl, DEPC-treated).

2xB&W (Binding&Washing) buffer (10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 2 M NaCl).

Dilution buffer (10 mM Tris-HCl (pH 7.5), 200 mM NaCl, 1 mM EDTA).

Protocol

1. RNA input



- The following protocol efficiently depletes ribosomal RNA from 100 ng up to 1.5 µg of total RNA (Fig. 1e). The procedure can be scaled up for higher RNA input.
2. Hybridization of biotinylated DNA oligonucleotides (40-mers) to ribosomal RNA
 - a) For oligonucleotide annealing the following reaction is set up:
 - 10 µl hybridization buffer
 - 10 µl RNA input (1 µg)
 - 1 µl of 100 µM biotinylated DNA probes
 - b) Gently mix the solution by pipetting and incubate at 68 °C for 10 min.
 - c) Immediately transfer the tubes to 37 °C for 30 min.
 3. Prepare Dynabeads MyOne streptavidin C1 (Invitrogen) according to the manufacturer's instruction as follows
 - a) For each sample use 120 µl (10 µg/µl) of bead slurry.
 - b) Wash the beads twice with an equal volume (or at least 1 ml) of Solution A. Add Solution A and incubate the mixture for 2 min. Then, place the tube on a magnet for 1 min and discard the supernatant.
 - c) Wash the beads once in Solution B. Split the washed beads into two separate tubes for two rounds of subtractive rRNA depletion (Round1 and Round2). Place the beads on a magnet for 1 min and discard Solution B.
 - d) Resuspend the beads for Round1 in 2xB&W buffer to a final concentration of 5 µg/µl (twice the original volume). The beads for Round1 will be used during the first round of rRNA depletion. For the second round of depletion, resuspend the beads for Round2 to a final concentration of 5 µg/µl in 1xB&W buffer. The beads for Round2 will be used during a second depletion step. Keep the beads at 37 °C until use.
 4. Capture of DNA-RNA hybrids using magnetic beads (step 2)
 - a) Briefly spin the tubes containing total RNA and probes. Then, add the following:
 - 100 µl dilution buffer.
 - 120 µl washed magnetic beads (5 µg/µl) in 2xB&W (Round1).
 Resuspend by pipetting up and down ten times. The final concentration of NaCl during this step is 1 M. Incubate the solution at 37 °C for 15 min. Gently mix the sample by occasional tapping.
 - b) Place on magnet for 2 min. Carefully remove the supernatant and add it to the additional 120 µl

of washed magnetic beads in 1xB&W (Round2). Incubate the mixture at 37 °C for 15 min with occasional gentle tapping.

- c) Place on magnet for 2 min. Carefully transfer the supernatant into a new tube and place on magnet for another 1 min to remove all traces of magnetic beads from the sample.
 - d) Transfer the supernatant into a fresh tube.
5. Use the RNA Clean & Concentrator-5 kit (Zymo research) to concentrate the ribodepleted samples, to carry out size selection and to digest any remaining DNA using DNase I treatment as described [27]

Ribosomal RNA depletion in planarian species related to *S. mediterranea*

Ribosomal DNA genes are among the most conserved sequences in all kingdoms of life. They are present in all organisms and are widely used for the construction of phylogenetic trees [28]. The latter is possible because of the low rate of nucleotide substitutions in rRNA sequences (about 1–2% substitutions occur per 50 million years based on bacterial 16S rRNA) [29]. The divergence of 18S rRNA sequence between different families of freshwater planarians lays in the range of 6–8%, while interspecies diversity does not exceed 4% [23]. Therefore, low rRNA divergence between taxa can be exploited for the design of universal probes for rRNA depletion in different organisms. To assess the specificity and universal applicability of our DNA probes, we depleted rRNA in flatworm species of the order Tricladida, all related to *S. mediterranea* (Fig. 2a). Total RNA separation profiles were analyzed before and after rRNA depletion of six planarian species from three different families. Two of these, *Dugesia japonica* and *Cura pinguis*, belong to the same family as *S. mediterranea*, the Dugesidae family. In addition, we examined three species from the family Planariidae (*Planaria torva*, *Polycelis nigra* and *Polycelis tenuis*) and one species from the genus *Camerata* of Uteriporidae (subfamily Uteriporinae). For all tested species our DNA probes proved efficient for the complete removal of rRNA, which migrated close to 2000 nts on all electropherograms (Fig. 2b). We note that the peak at about 100 nts in the rRNA-depleted samples represents a variety of small RNAs (5S and 5.8S rRNA, tRNAs, and other small RNA fragments) that evaded the size selection step aimed at retaining only fragments longer than 200 nts. Taken together, the probes developed for *S. mediterranea* can be utilized for the removal of ribosomal RNA in a multitude of planarian species and may even be generally applicable to all studied planarian species.

Comparison of RNA-seq libraries prepared by ribodepletion or poly(a) selection

To assess the efficiency of rRNA removal and the specificity of our DNA probes, we prepared and analyzed

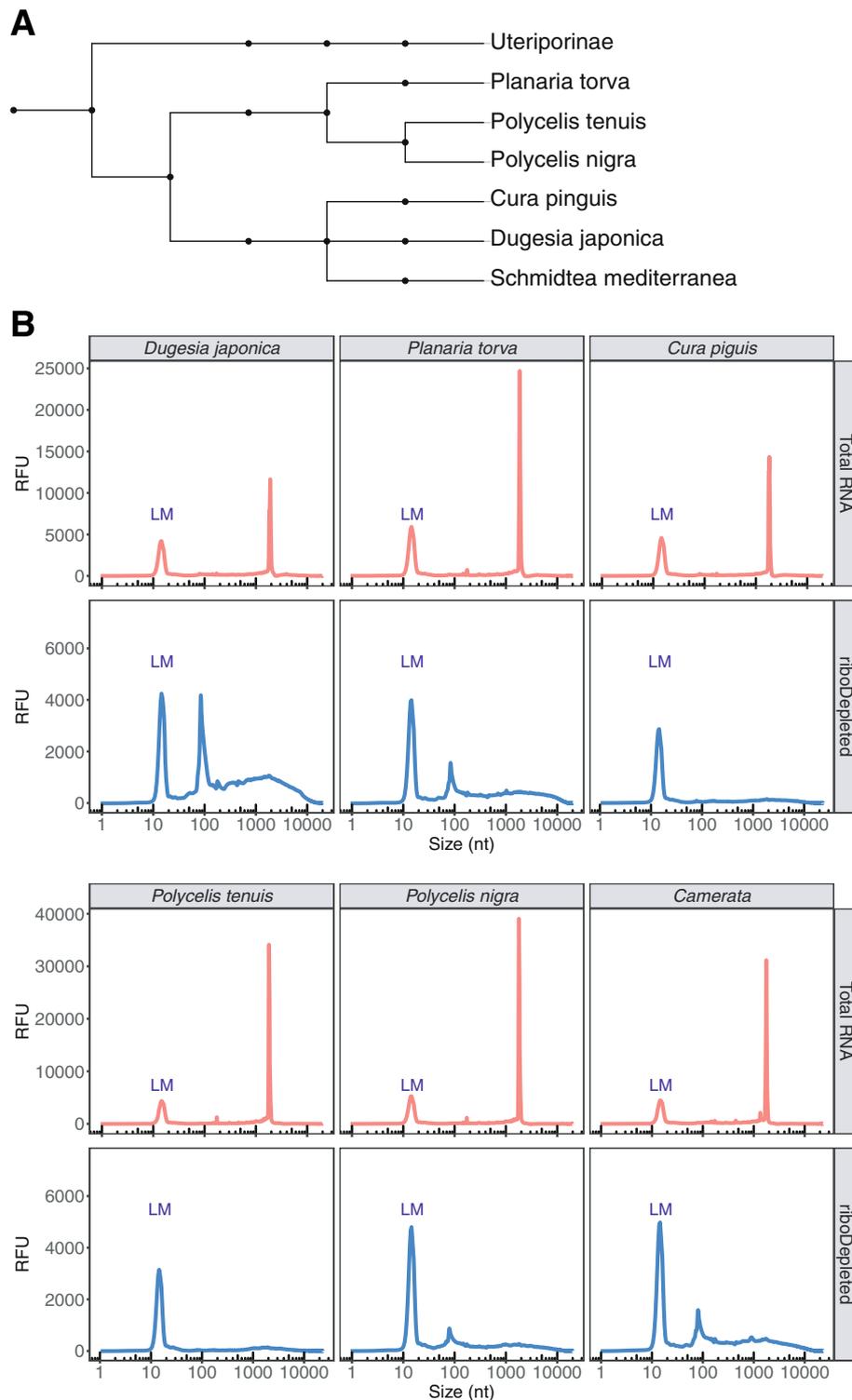


Fig. 2 Probes developed for *S. mediterranea* efficiently remove rRNA of other freshwater triclad. **a** Phylogenetic tree showing the taxonomic position of the analyzed planarian species. **b** Total RNA separation profile before and after rRNA depletion. In all species analyzed the 28S rRNA undergoes “gap deletion” maturation, which results in two co-migrating fragments. Both 28S fragments co-migrate with 18S rRNA, resulting in a single rRNA peak

RNA-seq libraries from ribodepleted total RNA from *S. mediterranea*. Total RNA was extracted from 100,000 FACS-sorted planarian neoblasts, resulting in 70–100 ng of input RNA. RNA-seq libraries were prepared and sequenced as described [27] following 15 cycles of PCR amplification. The subsequent analysis of sequenced libraries confirmed the efficient removal of rRNAs. Less than 2% of total sequenced reads constituted ribosomal RNA (Fig. 3a). Next, we compared our rRNA-depleted libraries with three publicly available planarian poly(A) enriched RNA-Seq datasets (poly(A) libraries) [30–32]. In case publicly available libraries were sequenced in paired-end mode, we analyzed only the first read of every pair to minimize the technical variation between libraries [33]. As shown in Fig. 3a, the ribodepleted libraries contained significantly less rRNA compared to all poly(A) enriched ones. Interestingly, the major rRNA species that remained after poly(A) selection was mitochondrial 16S rRNA (Fig. 3b). Although the planarian genome has a high A-T content (> 70%) [34], we could not attribute the overrepresentation of 16S rRNA in poly(A) libraries to a high frequency or longer stretches of A nucleotides as compared to other rRNA species (Fig. 3c). Moreover, using publicly available planarian poly(A)-position profiling by sequencing (3P-Seq) libraries [35], which allow the identification of 3'-ends of polyadenylated RNAs, no polyadenylation sites were detected in 16S rRNA. Therefore, we speculate that upon folding of 16S rRNA stretches of A nucleotides become exposed and facilitate the interaction with oligo-dT beads during transcript poly(A) selection.

We next assigned the analyzed datasets to the planarian genome. In ribodepleted libraries more than 13% of all mapped reads were assigned to intergenic regions, compared to 7–10.5% for poly(A)-enriched ones (Fig. 3d). In addition, the percentage of unmapped reads was higher in ribodepleted libraries and constituted about 17.6%, which is on average 2.4% more than in poly(A) datasets. We speculate that for ribodepleted libraries the proportion of reads mapping to intergenic regions will increase in the future, once complete assemblies of the planarian genome are available. Currently, the planarian genome assembly consists of 481 scaffolds [34]. To detect gene expression variabilities between the analyzed libraries, we performed principal component analysis for the clustering of gene expression data. Although all poly(A) selected libraries were grouped closer together along the PC1 scale, all four analyzed datasets appeared as separated clusters. This indicates considerable variation even amongst different batches of poly(A) libraries (Figs. 3e). One possible source of such variation might be the sequencing depth of the analyzed libraries, which varied considerably from 13 to 64 millions of mapped reads (Fig. 3f).

Next, to estimate the correlation between ribodepleted and poly(A) libraries, we calculated their Pearson correlation coefficients (Fig. 3g). We found the highest Pearson correlation between ribodepleted libraries and polyA B2 samples ($R = 0.94$, $p < 2.2e-16$) (Fig. 3f). This could be due to their similar sequencing depth compared to the other polyA libraries. The transcripts whose abundance was most significantly affected by poly(A) selection were found to be histone mRNAs that are known to lack polyA tails (Fig. 3g, h) [36]. Their expression level appeared to be 8–10 log₂ fold higher in our ribodepleted libraries. Moreover, in the ribodepleted libraries we also detected significantly higher expression levels for transposable elements (Fig. 3g, i). Out of 316 planarian transposable element families [37], 254 were on average upregulated 5.2, 3.5 and 4.0 log₂ fold as compared to polyA B1, polyA B2 and polyA B3 libraries, respectively (Fig. 3i). Moreover, the ribodepleted libraries revealed that Burro elements, giant retroelements found in planarian genome [34], gypsy retrotransposons, hAT and Mariner/Tc1 DNA transposons are the most active transposable elements in planarian stem cells. Although some transposable elements are polyadenylated, long-terminal repeat elements (LTRs) lack poly(A)-tails [38]. This renders their detection in poly(A)-enriched sample non-quantitative.

Non-specific depletion of coding transcripts in ribodepleted libraries

In using custom ribodepletion probes, our major concern was that the utilized probes would lead to unspecific co-depletion of planarian coding transcripts. To exclude this possibility, we first mapped our pool of 88 DNA probes in antisense orientation to the planarian transcriptome allowing up to 8 mismatches and gaps of up to 3 nts. This mapping strategy requires at least 75% of a DNA probe to anneal to its RNA target. It resulted in only 11 planarian genes to be potentially recognized by 20 DNA probes from our oligonucleotide pool. Next, we carried out a differential expression analysis of these 11 potentially targeted transcripts between the ribodepleted libraries and poly(A)-selected ones. The analysis revealed that 9 out of 11 potential targets were downregulated at least 1-fold in at least two poly(A) experiments (Fig. 4a). As the abundance of three transcripts (SMESG000014330.1 (rhodopsin-like orphan gpcr [39]), SMESG000068163.1 and SMESG000069530.1 (both without annotation)) was very low in all polyA libraries (< 0.6 transcripts per million (TPM)), we did not consider these any further. However, the remaining six transcripts were found to be significantly downregulated in ribodepleted libraries. For three of these targeted genes (SMESG000067473.1, SMESG000021061.1 and SMESG000044545.1) the probes map in regions that display significant RNA-seq coverage (Fig. 4b, Additional file 2: Figures S1a, S1b). Therefore, their

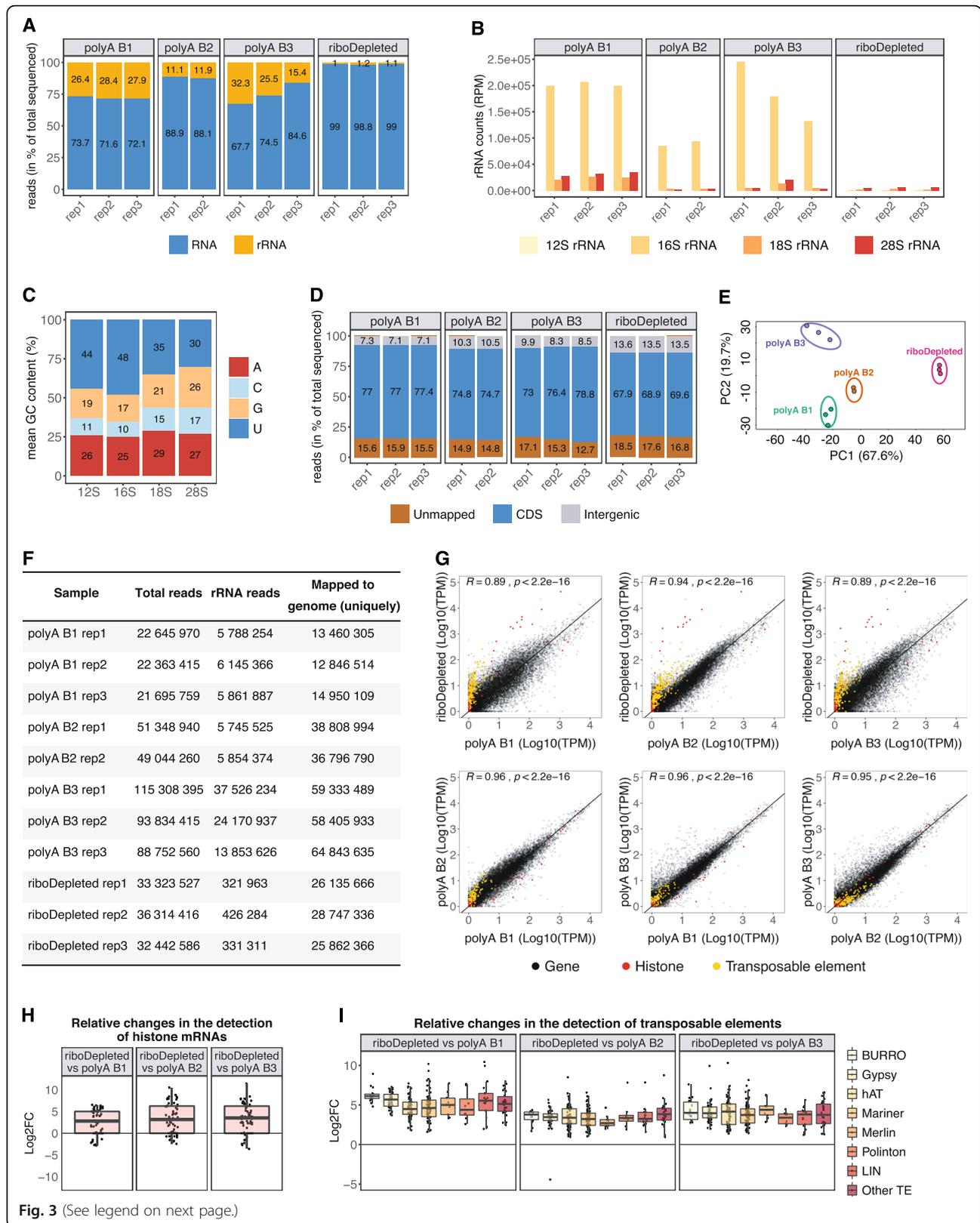


Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 Comparison of rRNA-depleted and poly(A)-enriched planarian RNA-seq libraries. **a** Percentage of rRNAs reads in the sequenced libraries prepared from rRNA-depleted or poly(A)-enriched RNA. **b** rRNAs species remaining in the final sequenced libraries. **c** Nucleotide content of planarian rRNA. **d** Percentage of sequenced reads mapped to coding (CDS) and intergenic regions in the planarian genome. **e** Principal component analysis (PCA) biplot of log₂ expression data for coding genes reveals distinct clustering of all analyzed RNA-seq experiments. **f** Sequencing depth and number of reads mapped to the planarian genome in analyzed ribodepleted and poly(A)-enriched samples. **g** Comparison of gene expression in transcripts per million (TPM) between planarian ribodepleted and poly(A)-enriched (polyA) RNA-Seq data. The Pearson's correlation coefficient is indicated. **h** Increased representation of histone mRNAs in ribodepleted libraries. **i** Boxplot of log₂ fold changes in the expression values of transposable elements between ribodepleted and poly(A)-enriched libraries

lower expression values in ribodepleted libraries is likely attributed to probe targeting. Intriguingly, for the remaining three targets (SMESG000066644.1, SMESG000043656.1 and SMESG000022863.1 annotated as RPL26 (ribosomal protein L26), COX11 (cytochrome c oxidase copper chaperone) and an unknown transcript, respectively) the probes were predicted to map to loci that do not exhibit RNA-seq coverage (Fig. 4c, Additional file 2: Figures S1C, S1D). The likely reason for this is inaccurate gene annotation. Alternatively, target regions might represent repetitive, multimapping sequences, which we excluded during read mapping. Taken together, our off-target analysis revealed that a maximum of 11 genes might be affected by our rRNA removal procedure - a very low number that underscores the specificity and efficiency of our depletion protocol.

Applicability of the described ribodepletion method to other organisms

To demonstrate the applicability of the developed rRNA workflow to other organisms, we employed our protocol to the depletion of ribosomal RNA from *Salmonella typhimurium* using a pool of organism-specific DNA probes (riboPOOL) developed by siTOOLS Biotech (Martinsried, Germany) (Fig. 5a). We compared the libraries resulting from the application of our newly developed procedure to the established rRNA depletion workflow that utilizes the Ribo-Zero rRNA Removal Kit (Bacteria) from Illumina. Removal of rRNA from a *S. typhimurium* sample using riboPOOL probes was as successful as a depletion reaction using Ribo-Zero, leaving as low as 3.4% rRNA in the final library (Fig. 5a). Moreover, an overall comparison of gene expression levels showed a high correlation (Pearson correlation $R = 0.98$, $p < 2.2 \times 10^{-16}$) between riboPOOL depleted libraries and libraries prepared with the Ribo-Zero kit (Fig. 5b). Taken together, the rRNA depletion workflow described in this manuscript is robust and easily applicable to any bacterial and eukaryotic species of choice utilizing organism-specific probes.

Discussion

For samples from typical model organisms, such as human, mouse and rat, there are numerous commercial kits available for the removal of rRNA, e.g. NEBNext from New England Biolabs, RiboGone from Takara and RiboCop from Lexogen. This also applies to typical

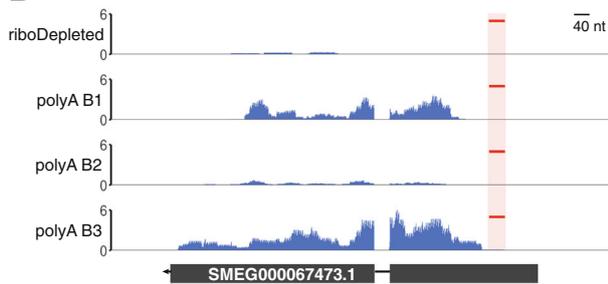
gram-positive and gram-negative bacteria (MICROBExpress from ThermoFisher and Ribominus from Invitrogen). Moreover, these kits can be utilized with a certain degree of compatibility for the depletion of rRNA in organisms of distinct phylogenetic groups (e.g. RiboMinus Eukaryote Kit for RNA-Seq, Invitrogen). However, as the breadth of molecularly tractable organisms has increased in the past decade, the necessity to develop organism-specific rRNA depletion techniques has risen as well [40–42]. To date, custom protocols either use biotinylated antisense probes along with streptavidin-coated magnetic beads for rRNA removal or rely on the digestion of DNA-RNA hybrids with RNase H [14, 43–45].

In this study, we describe a novel rRNA depletion workflow for the planarian flatworm *S. mediterranea*. Our protocol is based on the hybridization of biotinylated DNA probes to planarian rRNA followed by the subsequent removal of the resulting rRNA-DNA hybrids by using streptavidin-labeled magnetic beads. We tested the efficiency and specificity of our protocol by depleting rRNA from total RNA of neoblasts, planarian adult stem cells. A comparative analysis between ribodepleted and poly(A)-selected libraries revealed that our protocol retains all information present in poly(A) selected libraries. Over and above, we found ribodepleted libraries to contain additional information on histone mRNAs and transposable elements. The abundance of histone mRNAs in neoblasts is not unexpected, as planarian neoblasts are the only dividing cells in adult animals and thus require histones for packaging newly synthesized DNA [46, 47]. The high expression values of transposable elements likely reflects our ability to detect both non-poly(A) transcripts and degradation products of transposable elements generated by PIWI proteins loaded with transposon-specific piRNAs [48, 49]. Planarian PIWI proteins and their co-bound piRNAs are abundant in neoblasts and essential for planarian regeneration and animal homeostasis [15, 48–50]. Using our rRNA depletion protocol, we are now able to estimate the actual abundance of transposons and other repeats in planarians. This is important as these transcripts are generated from a large fraction of planarian genome (about 62% of the planarian genome comprise repeats and transposable elements) [34]. In addition, the planarian PIWI protein SMEDWI-3 is also involved in the degradation of multiple protein-coding transcripts in

A

Gene_id	TPM in polyA B1	TPM in polyA B2	TPM in polyA B3	TPM in ridoDepleted	LFC polyA+ B1	LFC polyA+ B2	LFC polyA+ B3	Annotation
SMESG000014330.1	0.27	0.24	0.38	0.005	-5,82	-5,30	-6,41	rhodopsin-like orphan gpcr
SMESG000068163.1	0.38	0.15	0.08	0.006	-6,13	-4,39	-3,89	NA
SMESG000069530.1	0.26	0.102	0.027	0.006	-5,99	-4,34	-2,51	NA
SMESG000022863.1	10.05	2.78	10.47	0.60	-4,23	FALSE	-4,34	NA
SMESG000021061.1	10.94	10.32	15.38	5.38	-1,27	-0,82	-1,77	MS3_11207 (MYDGF)
SMESG000067473.1	17.62	6.69	20.57	1.13	-4,21	-2,44	-4,45	MS3_08805 (GID8)
SMESG000044545.1	30.91	32.22	34.31	16.72	-1,11	-0,82	-1,25	REEP2
SMESG000043656.1	48.2	24.8	36.5	12.3	-2,20	-0,88	-1,81	COX11
SMESG000066644.1	1020.6	605.8	798.9	51.9	-4,53	-3,41	-4,19	RPL26

B



C

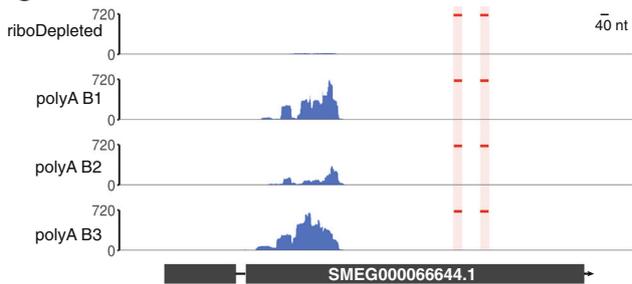
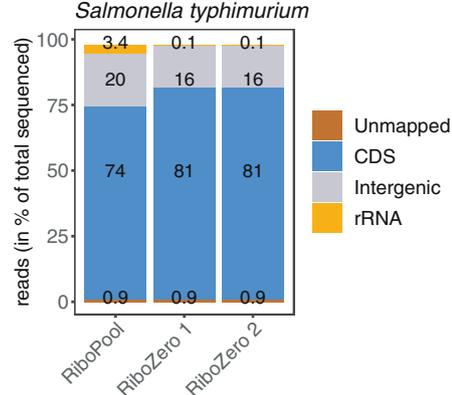


Fig. 4 Off-target analysis of DNA probes used for rRNA depletion. **a** Expression levels in TPM (transcripts per million) of nine transcripts targeted by probes utilized for ribodepletion. LFC denotes the log2 fold difference in the expression level of individual transcripts between ribodepleted and poly(A) enriched libraries. **b** RNA-seq coverage profile for SMESG000067473.1 in rRNA-depleted (riboDepleted) and poly(A) enriched (polyA B1, polyA B2, polyA B3) libraries. The location of antisense probes mapping to the transcripts is marked in red. **c** The same as in (B) for SMESG000066644.1

A



B

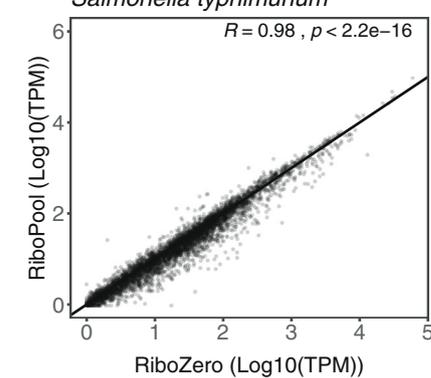


Fig. 5 Application of the developed rRNA workflow to other species using organism-specific probes. **a** Percentage of rRNA in sequenced libraries from *Salmonella typhimurium*. Libraries were prepared using our developed rRNA depletion workflow with organism-specific riboPOOL probes (siTOOLS Biotech) or the commercially available Ribo-Zero kit (Illumina). **b** Scatter plot comparing transcript abundance (TPM) between ribodepleted libraries using our developed workflow and the commercial Ribo-Zero kit. The Pearson's correlation coefficient is indicated

neoblasts [49]. Such mRNA degradation processes complicate the analysis of mRNA turnover using poly(A) enriched libraries, as these only represent mRNA steady-state levels. To study dynamic changes in mRNA levels is especially intriguing during neoblast differentiation, as then the steady-state levels of numerous mRNAs are changing [51, 52]. Using our rRNA-depletion protocol, we can now determine whether mRNA expression changes are due to altered transcription rates or due to increased degradation. Taken together, ribodepleted RNA-seq libraries are particularly valuable for the investigation of the piRNA pathway and RNA degradation processes as they retain the dynamics inherent to cellular RNA metabolism. Furthermore, by successfully depleting rRNA from other freshwater triclad species, we could demonstrate the versatility of the DNA probes designed for *S. mediterranea*. Last, we validated the efficiency of the developed workflow by removal of rRNA in the gram-negative bacterium *S. typhimurium*. Therefore, the proposed workflow likely serves as an efficient and cost-effective method for rRNA depletion in any organism of interest.

Conclusions

This study describes an rRNA depletion workflow for the planarian model system *S. mediterranea* and related freshwater triclads. It is based on the hybridization of 40-mer biotinylated DNA oligos to ribosomal RNA followed by the subtraction of formed DNA-RNA hybrids. The protocol is very robust and ensures the efficient removal of rRNA even from low input total RNA. Moreover, we suggest the general applicability of the presented workflow to any prokaryotic or eukaryotic organisms by using organism-specific pools of probes.

Materials and methods

Ribosomal RNA depletion

Ribosomal RNA depletion was conducted as described in the result section. To evaluate Fragment analyzer separation profiles, planarian total RNA (1000 ng for each sample) was subjected to rRNA depletion using varying concentrations of NaCl (0 mM, 50 mM, 250 mM, 500 mM) in the hybridization buffer.

Phylogenetic tree

The phylogenetic tree was constructed using NCBI taxonomic names at phyloT (<https://phyloT.biobyte.de>). The tree was visualized using the Interactive Tree of Life (iTOL) tool [53].

Processing of RNA-Seq libraries

Planarian RNA-seq data were processed as follows: Reads after removal of 3'-adapters and quality filtering with Trimmomatic (0.36) [54] were trimmed to a length of 50 nts. For libraries sequenced in pair-end mode, only

the first read of a pair was considered for the analysis. Next, sequences mapped to planarian rRNAs were removed with SortMeRNA [55]. Reads were assigned to the reference genome version SMESG.1 [34] or consensus transposable element sequences [37] in strand-specific mode. The abundance of transcripts was quantified with kallisto [56] using the settings: "--single -l 350 -s 30 -b 30". Differential gene expression analysis was performed with DeSeq2 [57]. To annotate RNA-Seq reads to coding regions (CDS), reads were mapped to the planarian genome using STAR [58] with the following settings: "--quantMode TranscriptomeSam --outFilterMultimapNmax 1.

RNA sequencing data from *Salmonella typhimurium* SL1344 were processed with READemption 0.4.3 using default parameters [59]. Sequenced reads were mapped to the RefSeq genome version NC_016810.1 and plasmids NC_017718.1, NC_017719.1, NC_017720.1.

Analysis of DNA probe specificity

DNA probe sequences were mapped to the planarian transcriptome SMEST.1 [60] using the BURST aligner (v0.99.7LL; DB15) [61] with the following settings "--fr -i .80 -m FORAGE". Only sequences that mapped to genes in antisense orientation with no more than 8 mismatches were considered as potential probe targets.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12864-019-6292-y>.

Additional file 1: Sequences of DNA oligonucleotides used for the removal of planarian rRNA. Assembled planarian rRNA sequences are indicated.

Additional file 2: (A)-(D) RNA-seq coverage profile for genes potentially targeted by ribodepletion probes in rRNA depleted (ribodepleted) and poly(A) enriched (polyA B1, polyA B2, polyA B3) libraries. The position of antisense probes mapping to the transcripts is marked in red.

Abbreviations

ITS: Internal transcribed spacer; LTR: Long terminal repeat; nts: Nucleotides; poly(A): Polyadenylated; rRNA: Ribosomal RNA; rRNA-depleted: Ribodepleted

Acknowledgements

We thank Jochen Rink, Mario Ivankovic and Miquel Vila Farré from the Max Planck Institute of Molecular Cell Biology and Genetics in Dresden for generously providing various flatworm species. We thank Jens Hör from the Institute for Molecular Infection Biology, University of Würzburg, for providing *Salmonella* RNA and IMGM Laboratories for library preparation and sequencing of *Salmonella* samples. We also acknowledge the KeyLab Genomics & Bioinformatics at the University of Bayreuth for Fragment Analyzer (Agilent) measurements.

Consent to publish

Not applicable.

Authors' contributions

I.V.K. and C.D.K. conceived and designed the study; I.V.K. and K.D. acquired data; E.J.R. assembled planarian rRNA sequences; I.V.K., E.J.R. and S.D. performed computation analyses; A.S.A. and C.D.K. supervised the study; I.V.K. and C.D.K. drafted the manuscript with input from all authors.

Funding

This work was supported by the Elite Network of Bavaria, the University of Bayreuth, the Paul Ehrlich and Ludwig Darmstaedter Prize for Young Researchers (to C.D.K) and the IZKF at the University of Würzburg (project Z-6). A.S.A. is an investigator of the Howard Hughes Medical Institute and the Stowers Institute for Medical Research.

Availability of data and materials

1. Planarian rRNA-depleted RNA-Seq datasets

Raw sequencing reads for planarian rRNA-depleted datasets were downloaded from the project GSE122199 (GSM3460490, GSM3460491, GSM3460492). The libraries were prepared as described [49]. Briefly, planarian rRNA depleted RNA-Seq libraries were prepared from 100,000 FACS-sorted planarian X1 cells as described [27] and sequenced on an Illumina Next-Seq 500 platform (single-end, 75 bp).

2. Publicly available RNA-Seq datasets

Raw sequencing reads for all datasets were downloaded from the Sequence read archive (SRA). Planarian polyA B1 rep1, polyA B1 rep2, polyA B1 rep3 correspond to SRR2407875, SRR2407876, and SRR2407877, respectively, from the Bioproject PRJNA296017 (GEO: GSE73027) [30]. Planarian polyA B2 rep1, polyA B2 rep2 samples correspond to SRR4068859, SRR4068860 from the Bioproject PRJNA338115 [32]. Planarian polyA B3 rep1, polyA B3 rep2, polyA B3 rep3 correspond to SRR7070906, SRR7070907, SRR7070908, respectively, (PRJNA397855) [31]. Only first read of the pair was analyzed for polyA B2 and polyA B3 from Bioprojects PRJNA338115 and PRJNA397855.

3. *Salmonella typhimurium* SL1344 RNA Seq datasets

In total, four samples were sequenced for *Salmonella typhimurium* SL1344 by IMG/M Laboratories GmbH (Martinsried, Germany) on an Illumina NextSeq 500 platform (single-end, 75 bp). One sample represented untreated total RNA, two samples comprised RiboZero and one RiboPool-treated total RNA. Sequencing data are available at the NCBI Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo>) under the accession number GSE132630.

Ethics approval and consent to participate

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Gene regulation by Non-coding RNA, Elite Network of Bavaria and University of Bayreuth, Universitätsstrasse 30, 95447 Bayreuth, Germany.

²Stowers Institute for Medical Research, 1000 East 50th Street, Kansas City, MO 64110, USA.

³Howard Hughes Medical Institute, Stowers Institute for Medical Research, 1000 East 50th Street, Kansas City, MO 64110, USA. ⁴Core Unit Systems Medicine, Institute for Molecular Infection Biology, University of Würzburg, Josef-Schneider-Str. 2, 97080 Würzburg, Germany.

Received: 18 June 2019 Accepted: 14 November 2019

Published online: 29 November 2019

References

- Elliott SA, Sánchez Alvarado A. The history and enduring contributions of planarians to the study of animal regeneration. *Wiley Interdiscip Rev Dev Biol.* 2013;2:301–26. <https://doi.org/10.1002/wdev.82>.
- Zeng A, Li H, Guo L, Gao X, McKinney S, Wang Y, et al. Prospectively isolated tetraspanin+ neoblasts are adult pluripotent stem cells underlying planaria regeneration. *Cell.* 2018;173:1593–608. <https://doi.org/10.1016/j.cell.2018.05.006>.
- Sahu S, Dattani A, Aboobaker AA. Secrets from immortal worms: what can we learn about biological ageing from the planarian model system? *Semin Cell Dev Biol.* 2017;70:108–21. <https://doi.org/10.1016/j.semcdb.2017.08.028>.
- Collins JJ. *Platyhelminthes.* *Curr Biol.* 2017;27:R252–6. <https://doi.org/10.1016/j.CUB.2017.02.016>.
- Egger B, Gschwentner R, Rieger R. Free-living flatworms under the knife: past and present. *Dev Genes Evol.* 2007;217:89–104. <https://doi.org/10.1007/s00427-006-0120-5>.
- Sikes JM, Newmark PA. Restoration of anterior regeneration in a planarian with limited regenerative ability. *Nature.* 2013;500:77–80. <https://doi.org/10.1038/nature12403>.
- Zhao W, He X, Hoadley KA, Parker JS, Hayes D, Perou CM. Comparison of RNA-Seq by poly (a) capture, ribosomal RNA depletion, and DNA microarray for expression profiling. *BMC Genomics.* 2014;15:419. <https://doi.org/10.1186/1471-2164-15-419>.
- Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, et al. Antisense transcription in the mammalian transcriptome. *Science.* 2005;309:1564–6. <https://doi.org/10.1126/science.1112009>.
- Cheng J, Kapranov P, Drenkow J, Dike S, Brubaker S, Patel S, et al. Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science.* 2005;308:1149–54. <https://doi.org/10.1126/SCIENCE.1108625>.
- Kim T-K, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, et al. Widespread transcription at neuronal activity-regulated enhancers. *Nature.* 2010;465:182. <https://doi.org/10.1038/NATURE09033>.
- Finnegan DJ. Retrotransposons. *Curr Biol.* 2012;22:R432–7. <https://doi.org/10.1016/j.CUB.2012.04.025>.
- Reuter M, Berninger P, Chuma S, Shah H, Hosokawa M, Funaya C, et al. Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. *Nature.* 2011;480:264–7. <https://doi.org/10.1038/nature10672>.
- Wang W, Yoshikawa M, Han BW, Izumi N, Tomari Y, Weng Z, et al. The initial uridine of primary piRNAs does not create the tenth adenine that is the hallmark of secondary piRNAs. *Mol Cell.* 2014;56:708–16. <https://doi.org/10.1016/j.molcel.2014.10.016>.
- Kukutla P, Steritz M, Xu J. Depletion of ribosomal RNA for mosquito gut metagenomic RNA-seq. *J Vis Exp.* 2013. <https://doi.org/10.3791/50093>.
- Reddian PW, Oviedo NJ, Jennings JR, Jenkin JC, Sánchez Alvarado A. SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. *Science.* 2005;310:1327–30. <https://doi.org/10.1126/science.1111610>.
- Hayashi T, Asami M, Higuchi S, Shibata N, Agata K. Isolation of planarian X-ray-sensitive stem cells by fluorescence-activated cell sorting. *Develop Growth Differ.* 2006;48:371–80. <https://doi.org/10.1111/j.1440-169X.2006.00876.x>.
- Wetmur JG. DNA probes: applications of the principles of nucleic acid hybridization. *Crit Rev Biochem Mol Biol.* 1991;26:227–59. <https://doi.org/10.3109/10409239109114069>.
- Green MR, Sambrook J. *Molecular cloning, 3-volume set : a laboratory manual.* 2012. ISBN 978-1-936113-42-2.
- Sun S, Xie H, Sun Y, Song J, Li Z. Molecular characterization of gap region in 28S rRNA molecules in brine shrimp *Artemia parthenogenetica* and planarian *Dugesia japonica*. *Biochem.* 2012;77:411–7. <https://doi.org/10.1134/S000629791204013X>.
- van Keulen H, Mertz PM, LoVerde PT, Shi H, Rekosch DM. Characterization of a 54-nucleotide gap region in the 28S rRNA gene of *Schistosoma mansoni*. *Mol Biochem Parasitol.* 1991;45:205–14. [https://doi.org/10.1016/0166-6851\(91\)90087-M](https://doi.org/10.1016/0166-6851(91)90087-M).
- Ware VC, Renkawitz R, Gerbi SA. rRNA processing: removal of only nineteen bases at the gap between 28S alpha and 28S beta rRNAs in *Sciara coprophila*. *Nucleic Acids Res.* 1985;13:3581–97. <https://doi.org/10.1093/nar/13.10.3581>.
- Carranza S, Bagnà J, Riutort M. Origin and evolution of Paralogous rRNA gene clusters within the flatworm family Dugesidae (Platyhelminthes, Tricladida). *J Mol Evol.* 1999;49:250–9. <https://doi.org/10.1007/PL00006547>.
- Carranza S, Giribet G, Ribera C, Riutort M. Evidence that two types of 18S rDNA coexist in the genome of *Dugesia* (Schmidtea) mediterranea (Platyhelminthes, Turbellaria, Tricladida). *Mol Biol Evol.* 1996;13:824–32. <https://doi.org/10.1093/oxfordjournals.molbev.a025643>.
- Draper DE. A guide to ions and RNA structure. *RNA.* 2004;10:335–43. <https://doi.org/10.1261/RNA.5205404>.
- Lambert D, Leipply D, Shiman R, Draper DE. The influence of monovalent cation size on the stability of RNA tertiary structures. *J Mol Biol.* 2009;390:791–804. <https://doi.org/10.1016/j.jmb.2009.04.083>.
- Garrey SM, Blech M, Riffell JL, Hankins JS, Stickney LM, Diver M, et al. Substrate binding and active site residues in RNases E and G: role of the 5'-sensor. *J Biol Chem.* 2009;284:31843–50. <https://doi.org/10.1074/jbc.M109.063263>.
- Zhang Z, Theurkauf WE, Weng Z, Zamore PD. Strand-specific libraries for high throughput RNA sequencing (RNA-Seq) prepared without poly(a) selection. *Silence.* 2012;3:9. <https://doi.org/10.1186/1758-907X-3-9>.
- Woese CR. Interpreting the universal phylogenetic tree. *Proc Natl Acad Sci U S A.* 2000;97:8392–6. <https://doi.org/10.1073/pnas.97.15.8392>.
- Ochman H, Elwyn S, Moran NA. Calibrating bacterial evolution. *Proc Natl Acad Sci.* 1999;96:12638–43. <https://doi.org/10.1073/pnas.96.22.12638>.

30. Duncan EM, Chitsazan AD, Seidel CW, Alvarado AS. Erratum: Set1 and MLL1/2 target distinct sets of functionally different genomic loci in vivo. *Cell Rep*. 2016;17:930. <https://doi.org/10.1016/j.celrep.2016.09.071>.
31. Schmidt D, Reuter H, Hüttner K, Ruhe L, Rabert F, Seebeck F, et al. The integrator complex regulates differential snRNA processing and fate of adult stem cells in the highly regenerative planarian *Schmidtea mediterranea*. *PLoS Genet*. 2018;14:e1007828. <https://doi.org/10.1371/journal.pgen.1007828>.
32. Mihaylova Y, Abnave P, Kao D, Hughes S, Lai A, Jaber-Hijazi F, et al. Conservation of epigenetic regulation by the MLL3/4 tumour suppressor in planarian pluripotent stem cells. *Nat Commun*. 2018;9:3633. <https://doi.org/10.1038/s41467-018-06092-6>.
33. Williams AG, Thomas S, Wyman SK, Holloway AK. RNA-seq data: challenges in and recommendations for experimental design and analysis. *Curr Protoc Hum Genet*. 2014;83:11.13.1–20. <https://doi.org/10.1002/0471142905.hg1113s83>.
34. Grohme MA, Schloissnig S, Rozanski A, Pippel M, Young GR, Winkler S, et al. The genome of *Schmidtea mediterranea* and the evolution of core cellular mechanisms. *Nature*. 2018;554:56–61. <https://doi.org/10.1038/nature25473>.
35. Lakshmanan V, Bansal D, Kulkarni J, Poduval D, Krishna S, Sasidharan V, et al. Genome-wide analysis of polyadenylation events in *Schmidtea mediterranea*. *G3 (Bethesda)*. 2016;6(10):3035–48. <https://doi.org/10.1534/g3.116.031120>.
36. Marzluff WF, Wagner EJ, Duronio RJ. Metabolism and regulation of canonical histone mRNAs: life without a poly(a) tail. *Nat Rev Genet*. 2008;9:843–54. <https://doi.org/10.1038/nrg2438>.
37. Bao W, Kojima KK, Kohany O. Repbase update, a database of repetitive elements in eukaryotic genomes. *Mob DNA*. 2015;6:11. <https://doi.org/10.1186/s13100-015-0041-9>.
38. Kazazian HH. Mobile elements : drivers of genome evolution. *Science*. 2004;303:1626–32. <https://doi.org/10.1126/science.1089670>.
39. Saberi A, Jamal A, Beets I, Schoofs L, Newmark PA. GPCRs direct Germline development and somatic gonad function in planarians. *PLoS Biol*. 2016;14:e1002457. <https://doi.org/10.1371/journal.pbio.1002457>.
40. Dietrich MR, Ankeny RA, Chen PM. Publication trends in model organism research. *Genetics*. 2014;198:787–94. <https://doi.org/10.1534/genetics.114.169714>.
41. Goldstein B, King N. The future of cell biology: emerging model organisms. *Trends Cell Biol*. 2016;26:818–24. <https://doi.org/10.1016/j.tcb.2016.08.005>.
42. Russell JJ, Theriot JA, Sood P, Marshall WF, Landweber LF, Fritz-Laylin L, et al. Non-model model organisms. *BMC Biol*. 2017;15:55. <https://doi.org/10.1186/s12915-017-0391-5>.
43. O'Neil D, Glowatz H, Schlumpberger M. Ribosomal RNA depletion for efficient use of RNA-Seq capacity. In: Ausubel FM, editor. *Current protocols in molecular biology*. Hoboken: Wiley; 2013. p. 4.19.1–8. <https://doi.org/10.1002/0471142727.mb0419s103>.
44. Li S-K, Zhou J-W, Yim A-K-Y, Leung A-K-Y, Tsui S-K-W, Chan T-F, et al. Organism-specific rRNA capture system for application in next-generation sequencing. *PLoS One*. 2013;8:e74286. <https://doi.org/10.1371/journal.pone.0074286>.
45. Morlan JD, Qu K, Sinicropi DV. Selective depletion of rRNA enables whole transcriptome profiling of archival fixed tissue. *PLoS One*. 2012;7:e42882. <https://doi.org/10.1371/journal.pone.0042882>.
46. Marzluff WF, Koreski KP. Birth and death of histone mRNAs. *Trends Genet*. 2017;33:745–59. <https://doi.org/10.1016/j.tig.2017.07.014>.
47. Reddien PW. Specialized progenitors and regeneration. *Development*. 2013;140:951–7. <https://doi.org/10.1242/dev.080499>.
48. Shibata N, Kashima M, Ishiko T, Nishimura O, Rouhana L, Misaki K, et al. Inheritance of a nuclear PIWI from pluripotent stem cells by somatic descendants ensures differentiation by silencing transposons in planarians. *Dev Cell*. 2016;37:226–37. <https://doi.org/10.1016/j.devcel.2016.04.009>.
49. Kim IV, Duncan EM, Ross EJ, Gorbovytska V, Nowotarski S, Elliott SA, et al. Planarians recruit piRNAs for mRNA turnover in adult stem cells. *Genes Dev*. 2019. <https://doi.org/10.1101/gad.322776.118>.
50. Palakodeti D, Smielewska M, Lu YC, Yeo GW, Graveley BR. The PIWI proteins SMEDWI-2 and SMEDWI-3 are required for stem cell function and piRNA expression in planarians. *RNA*. 2008;14:1174–86. <https://doi.org/10.1261/ma.1085008>.
51. Solana J, Gamberi C, Mihaylova Y, Grosswendt S, Chen C, Lasko P, et al. The CCR4-NOT complex mediates Deadenylation and degradation of stem cell mRNAs and promotes planarian stem cell differentiation. *PLoS Genet*. 2013;9:e1004003. <https://doi.org/10.1371/journal.pgen.1004003>.
52. Lou C-H, Shum EY, Wilkinson MF. RNA degradation drives stem cell differentiation. *EMBO J*. 2015;34:1606–8. <https://doi.org/10.15252/embj.201591631>.
53. Letunic I, Bork P. Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res*. 2019. <https://doi.org/10.1093/nar/gkz239>.
54. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114–20. <https://doi.org/10.1093/bioinformatics/btu170>.
55. Kopylova E, Noé L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics*. 2012;28:3211–7. <https://doi.org/10.1093/bioinformatics/bts611>.
56. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol*. 2016;34:525–7. <https://doi.org/10.1038/nbt.3519>.
57. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550. <https://doi.org/10.1186/s13059-014-0550-8>.
58. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29:15–21. <https://doi.org/10.1093/bioinformatics/bts635>.
59. Forstner KU, Vogel J, Sharma CM. READemption—a tool for the computational analysis of deep-sequencing-based transcriptome data. *Bioinformatics*. 2014;30:3421–3. <https://doi.org/10.1093/bioinformatics/btu533>.
60. Rozanski A, Moon H, Brandl H, Martín-Durán JM, Grohme MA, Hüttner K, et al. PlanMine 3.0-improvements to a mineable resource of flatworm biology and biodiversity. *Nucleic Acids Res*. 2019;47:D812–20. <https://doi.org/10.1093/nar/gky1070>.
61. Al-Ghalith, Knights G, Knights D. BURST enables optimal exhaustive DNA alignment for big data. 2017. <https://doi.org/10.5281/zenodo.806850>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

