MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2012) **12**, 312–322 doi: 10.1111/j.1755-0998.2011.03097.x

Optimization of preservation and storage time of sponge tissues to obtain quality mRNA for next-generation sequencing

ANA RIESGO,* ALICIA R. PÉREZ-PORRO,*† SUSANA CARMONA,‡ SALLY P. LEYS§ and GONZALO GIRIBET*

*Museum of Comparative Zoology, Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA, †Center for Advanced Studies of Blanes, c/Accés a la Cala St. Francesc, 14, 17300 Blanes, Girona, Spain, ‡Harvard Social Cognitive Neuroscience Lab, 52 Oxford Street, Cambridge, MA 02138, USA, §Department of Biological Sciences, University of Alberta, T6G 2E9, Edmonton, Alberta, Canada

Abstract

Transcriptome sequencing with next-generation sequencing technologies has the potential for addressing many long-standing questions about the biology of sponges. Transcriptome sequence quality depends on good cDNA libraries, which requires high-quality mRNA. Standard protocols for preserving and isolating mRNA often require optimization for unusual tissue types. Our aim was assessing the efficiency of two preservation modes, (i) flash freezing with liquid nitrogen (LN_2) and (ii) immersion in RNA*later*, for the recovery of high-quality mRNA from sponge tissues. We also tested whether the long-term storage of samples at −80 °C affects the quantity and quality of mRNA. We extracted mRNA from nine sponge species and analysed the quantity and quality (A260/230 and A260/280 ratios) of mRNA according to preservation method, storage time, and taxonomy. The quantity and quality of mRNA depended significantly on the preservation method used (LN_2 outperforming RNA*later*), the sponge species, and the interaction between them. When the preservation was analysed in combination with either storage time or species, the quantity and A260/230 ratio were both significantly higher for LN_2-preserved samples. Interestingly, individual comparisons for each preservation method over time indicated that both methods performed equally efficiently during the first month, but RNA*later* lost efficiency in storage times longer than 2 months compared with flash-frozen samples. In summary, we find that for long-term preservation of samples, flash freezing is the preferred method. If LN_2 is not available, RNA*later* can be used, but mRNA extraction during the first month of storage is advised.

Keywords: flash freezing, illumina, messenger RNA, nucleic acids, RNA isolation, RNA*later*

Received 13 June 2011; revision received 19 October 2011; accepted 25 October 2011

Introduction

Sponge genomics and transcriptomics are growing fields. So far, only the genome of *Amphimedon queenslandica* (Srivastava *et al.* 2010) and expressed sequence tags for a handful of sponges are publically available. These genomic resources have been explored for phylogenomic approaches (e.g. Dunn *et al.* 2008; Hejnol *et al.* 2009; Philippe *et al.* 2009) and/or gene discovery, developmental biology, novelty acquisition analysis and functional genetics (e.g. Nichols *et al.* 2006; Adamska *et al.* 2007, 2010; Larroux *et al.* 2007; Sullivan *et al.* 2008; Funayama *et al.* 2010; Harcet *et al.* 2010; Rivera *et al.* 2011; Leys & Riesgo in press). With the current reduction in sequencing costs of next-generation sequencing (NGS) platforms (e.g. 454, Illumina, and SOLID), genomic and transcriptomic data are becoming broadly accessible.

Correspondence: Ana Riesgo, Fax: +1 617 495 5667;
E-mail: ariesgo@oeb.harvard.edu

Next-generation sequencing permits a great diversity of genetic approaches using both the DNA (genomic) and RNA (transcriptomic) perspectives. NGS is currently used for re-sequencing in genome projects, phylogenomics, gene expression profiling, small non-coding RNA profiling, protein-coding gene annotation, aberrant transcription events, mutation detection, single nucleotide polymorphism detection, etc. (Morozova & Marra 2008; Shendure & Ji 2008; Reis-Filho 2009). Transcriptomic data are more advantageous than genomic data in several respects. Transcriptomes are from 10 to 100 times smaller than genomes, and the new *de novo* assemblers (Zerbino & Birney 2008; Blow 2009; Simpson *et al.* 2009) make the

© 2011 Blackwell Publishing Ltd

handling of such amounts of sequencing data (from 10 to 250 million short reads) easier and faster than whole-genome data when no reference genome is available. Also, transcriptomic data simplify the identification of large numbers of orthologous genes across non-model species, facilitating not only the assembly of large data sets for phylogenomic analysis (e.g. Dunn *et al.* 2008; Hejnol *et al.* 2009; Philippe *et al.* 2009) but also more restricted studies of speciation among sympatric species (e.g. Elmer *et al.* 2010). Transcriptomic data can also be used to develop molecular markers for PCR amplification, study alternative splicing, create targeted sequencing assays, and study gene regulation and DNA-protein interactions (reviewed in Ekblom & Galindo 2011). Another advantage of transcriptomic data is that it also further contains information on the level of gene expression, which has led to an emergent area of interest both for developmental biologists, who previously relied on microarray analysis for those data (e.g. Toth *et al.* 2007; Marioni *et al.* 2008; Sultan *et al.* 2008; Vera *et al.* 2008), and for researchers investigating animals for which microarrays were a long way from being available or optimized (e.g. Gilad *et al.* 2009).

Complementary DNA (cDNA) libraries used for sequencing transcriptomic data are obtained from pure mRNA extracted from tissues, and this is one of the greatest disadvantages of transcriptomic approaches. Compared to DNA, working with RNA can be technically difficult, especially when the RNA extractions are to be performed from unusual tissue types where protocols are often not optimized. Given that RNA is prone to rapid degradation, if samples are to be stored for any length of time (sometimes even just the time from collection to processing may be enough for degradation to begin), it is necessary to preserve tissues to prevent damage to nucleic acids. Recently, Simister *et al.* (2011) have shown that both liquid nitrogen and RNA*later* performed favourably when preserving the purity and integrity of total RNA in demosponge tissues, even though liquid nitrogen outperformed RNA*later*. However, only total RNA quality control was performed, and comparisons with mRNA isolation data are not available. Because cDNA synthesis relies on mRNA isolation, assessing the quality of starting mRNA is a key step in the process, and therefore, it is important to develop optimized protocols for preservation and storage of tissues.

Sponge transcriptomic projects can certainly take advantage of standardization of RNA preservation protocols to ensure the maximum recovery of mRNA. The aim of our study was twofold: 1 – assessing the best RNA preservation method for sponge tissues and 2 – testing whether long-term storage of samples could result in mRNA degradation. Here, we compared the quantity and quality (measured as A260/230 and A260/280 ratios) of mRNA extracted from nine different sponge species belonging to the four extant poriferan classes preserved either by flash freezing or by immersion in RNA*later*. To assess the effect of storage time on the quantity and quality of mRNA, we compared the results of extractions stored for periods <1 month to more than 12 months.

Methods

Species and sample collection

We collected samples from nine sponge species, belonging to the four extant poriferan classes (Table 1). *Aphrocallistes vastus* (Hexactinellida, Lyssacinosida) and *Asbestopluma occidentalis* (Demospongiae, Poecilosclerida) tissues were collected during a scientific cruise in the Fraser Ridge (Strait of Georgia, Vancouver Island, Canada) with the Canadian ROV ROPOS in 2009. *Spongilla lacustris* (Demospongiae, Haplosclerida) was collected at Frederick Lake (Vancouver Island, Canada) in 2011. *Sycon coactum* (Calcarea, Leucosolenida) was collected from ropes hanging off docks at the Bamfield Marine Sciences Centre (Bamfield, British Columbia, Canada) during 2009–2011. *Corticium candelabrum* (Homoscleromorpha, Homosclerophorida), *Crella elegans* (Demospongiae, Poecilosclerida), and *Petrosia ficiformis* (Demospongiae, Haplosclerida) were collected using SCUBA diving in the subtidal rocky coast of Northeastern Spain in various sampling trips in 2009 and 2010. *Lissodendoryx colombiensis* (Demospongiae, Poecilosclerida) was collected in a shallow reef at Isla Colón (Bocas del Toro, Panama) in 2010. *Cliona* sp. (Demospongiae, Hadromerida) was collected at Niles Channel (Summerland Key, FL, USA) in 2010. More details about collecting localities are available upon request.

Sample preparation

To avoid contaminations from epibionts, tissues were carefully cleaned using a stereomicroscope. Tissue excisions were always performed with sterilized razor blades rinsed with RNAseZap® (Ambion, TX, USA). All cleaning procedures were operated as quickly as possible to avoid RNA degeneration in an RNAse-free and cold environment (in dishes kept on ice, for example).

Preservation of tissues was performed the same day tissues were collected, usually 1–5 h later depending on the time required for epibiont removal, and up to 8 h in the cases of deep-sea species where gear may take up several hours to surface and samples may require longer periods of sorting. Tissues were cut into pieces from 0.25 to 0.5 cm in thickness. Usually, 80 mg of tissue were placed in each Eppendorf tube. Because of the large amount of spicules present in the body of *A. vastus*,

© 2011 Blackwell Publishing Ltd

Table 1 Number of extractions performed on each species preserved in LN_2/−80 °C and RNA*later* at each storage time: 0–1, 2–5, 6–11 and more than 12 months (12+)

Preservation	LN_2/−80 °C				RNA*later*				
Storage time (months)	0–1	2–5	6–11	12+	0–1	2–5	6–11	12+	Totals
Aphrocallistes vastus	–	1	–	4	–	–	–	5	10
Asbestopluma occidentalis	–	–	–	–	–	–	–	2	2
Cliona sp.	–	–	3	–	–	–	–	–	3
Corticium candelabrum	1	2	–	–	–	1	3	–	7
Crella elegans	**4**	**3**	–	–	3	**12**	1	3	26
Lissodendoryx colombiensis	–	–	–	–	–	–	2	–	2
Petrosia ficiformis	2	6	1	1	1	1	4	–	16
Spongilla lacustris	1	–	**3**	–	–	–	–	–	4
Sycon coactum	3	6	2	1	1	1	1	–	15
Totals	11	18	9	6	5	15	11	10	85

Bold numbers represent samples where one replicate was removed prior to subsequent analysis because of very low values.

P. ficiformis and *S. coactum*, tissue pieces were larger (150–200 mg) to make the amount of sponge cells comparable with the other species (spicules represented about 80% of the total weight for *A. vastus* and *S. coactum*, and up to 50–80% for *P. ficiformis*). Two different standard methods of RNA preservation were tested for sponge samples: (i) flash freezing in liquid nitrogen and immediate storage at −80 °C, and (ii) immersion of the tissues in at least 10 volumes of RNA*later*® at 4 °C for 1 h, overnight incubation at −20 °C, and subsequent storage in the same buffer at −80 °C until RNA extraction (sometimes samples placed in RNA*later* were transported back to the lab at room temperature, and then stored at −80 °C). We refer to these treatments as liquid nitrogen or flash freezing and RNA*later*, respectively.

To test whether short- or long-term storage of tissues affected the quantity or quality of the yield of mRNA, RNA extractions were performed between a few days to 2 years after collection and preservation (Table 1). We defined four categories as follows: 0–1, 2–5, 6–11 and longer than 12 months (12 months+).

mRNA extractions

Two different methods of RNA extraction were used: (i) total RNA extraction followed by mRNA purification, and (ii) direct mRNA extraction. For Total RNA extraction, we used a standard trizol-based method using Tri-Reagent® (Ambion) following the manufacturer's protocol. Subsequent mRNA purification was performed with the Dynabeads® mRNA Purification Kit for mRNA Purification from Total RNA preps following manufacturer's instructions. Details about these procedures are given in Data S1 (Supporting Information).

For the direct extraction of mRNA, we used two different kits: (i) The New England Biolabs Magnetic mRNA Isolation Kit (New England Biolabs, Ipswich, MA, USA), and (ii) the Dynabeads® mRNA DIRECT™ Kit (Invitrogen, Carlsbad, CA, USA). Manufacturer's instructions were followed in both cases, with minor modifications detailed in Data S1 (Supporting Information).

Quantity and quality control of mRNA

Quantity and quality (purity and integrity) of mRNA were assessed by two different methods. We measured the absorbance at different wavelengths using a NanoDrop ND-1000 UV spectrophotometer (Thermo Fisher Scientific, Wilmington, MA, USA). RNA has an absorption maximum at 260 nm and the ratio of the absorbance at 260 and 280 nm and 260 and 230 nm is used to assess the RNA purity. A 260/230 ratio was used to estimate the presence of contaminants such as salts, carbohydrates, peptides, etc., while A260/280 ratio was used to estimate the purity of mRNA. Both ratios should show values close to 2.0–2.2 for pure mRNA. Also, capillary electrophoresis in an RNA Pico 6000 chip was performed using an Agilent Bioanalyzer 2100 System with the 'mRNA pico Series II' assay (Agilent Technologies, CA, USA). Integrity of mRNA was estimated by the electropherogram profile and lack of rRNA contamination (based on rRNA peaks for 18S and 28S rRNA given by the BIOANALYZER software).

Statistical analyses

The original sample consisted of 85 extractions from nine species (Table 1). Only the samples with mRNA yield higher than 3.5 ng/μL were considered (n = 80), which resulted in discarding five samples of *C. elegans* and one sample of *S. lacustris*. We only used mRNA yields over 3.5 ng/μL because we consider lower yields as a failed

© 2011 Blackwell Publishing Ltd

extraction and also because no A260/230 and A260/280 ratios could be reliably estimated for those extractions.

We performed an independent sample *t*-test analysis by groups to analyse the effect of the preservation [RNA-*later* and liquid nitrogen (LN_2)] on the quantity (mRNA concentration measured in nanogram per microlitre) and quality (A260/230 and A260/280 ratios) of mRNA. Because the quantity of mRNA did not follow a normal distribution according to a Kolmogorov–Smirnov test (K–S d = 0.19852; $P < 0.01$), we logarithmically transformed the original values (K–S d = 0.11268; $P > 0.20$).

A one-way analysis of variance (ANOVA) was used to test the effect of the storage time (0–1, 2–5, 6–11 and 12 months or more) on the quantity and quality of mRNA. Likewise, a one-way ANOVA was used to test the effect of the nine studied species on the quantity and quality of mRNA.

To analyse whether the effect of the preservation on the quantity and quality of mRNA varied across species or through time, we only selected those species in which both preservation methods were used (*A. vastus*, *Corticium candelabrum*, *C. elegans*, *P. ficiformis*, and *S. coactum*; $n = 60$). We performed two by four ANOVAs to test the individual effects of each factor and also their interactions. These analyses were followed by independent sample *t*-tests to assess within each of the species or time periods which preservation yielded higher mRNA concentrations and/or quality values.

We also performed *post hoc* pair-wise comparisons using Tukey's test to explore whether there were differences in the quantity and quality of mRNA between the different species and between the different storage time periods for each preservation type.

All the statistical analyses were performed using STATISTICA 10.0 and the graphs plotted with SIGMAPLOT 10.0.

Next-generation sequencing

Next-generation sequencing was performed using Illumina GA, GAII, and HiSeq (Illumina Inc., CA, USA) instruments at the FAS Center for Systems Biology at Harvard University. mRNA concentrations between 1.5 and 0.091 μg were used for random primed first-strand synthesis using Superscript II and III (Invitrogen), followed by second-strand synthesis with DNA Polymerase I and enzymatic fragmentation using the NEBNext® dsDNA Fragmentase (New England BioLabs). End repair of the double stranded cDNA (ds cDNA) was performed with NEBNext® End Repair Module (New England BioLabs) and an additional dAMP was incorporated with the NEBNext® dA-Tailing Module (New England BioLabs). Homemade adapters were ligated to the ds cDNA fragments with the NEBNext® Quick Ligation Module (New England BioLabs). Size-selected cDNA fragments of around 300 base pairs cut from a 2% agarose gel were amplified using Illumina PCR Primers for Paired-End reads (Illumina Inc.) and 18 cycles of the PCR program 98 °C-30 s, 98 °C-10 s, 65 °C-30 s, 72 °C-30 s, followed by an extension step of 5 min at 72 °C. Fragments resulting from the PCR were sequenced with a read length of 100 or 150 bp.

We also used the TruSeq RNA Sample Prep Kit (Illumina Inc.) to prepare samples of *Cliona* sp., *Corticium candelabrum*, *P. ficiformis*, and *C. elegans* (see Table 2) following the manufacturer's instructions. Fragmentation was performed on mRNA for 1.5 min, and fragments of 350 bp were targeted.

The concentration of the cDNA libraries was measured with the QubiT® dsDNA High Sensitivity (HS) Assay Kit and the QubiT® Fluoremeter (Invitrogen). The quality of the library was checked by using the 'HS DNA assay' in a DNA chip for Agilent Bioanalyzer 2100 (Agilent Technologies). Four different profiles of cDNA libraries were obtained consistently: (1), a tight band of targeted size with high cDNA concentration; (2), a tight band of targeted size and additional 'bumps' of smaller or larger fragments; (3), no bands; (4), a tight band of targeted size with low cDNA concentration. cDNA libraries were considered successful when the final concentration was higher than 1 ng/μL and the profile was correct (1 or 4) (see Fig. S1, Supporting Information). Libraries showing a profile of type 2 were never sequenced, because they could have been the result of a PCR artefact and affect the efficiency of the NGS sequencing platform.

Successful libraries were brought to 10, 7 or 3 nM depending on the initial concentration, to run on the sequencing platforms, which were selected upon availability.

Thinning and trimming analyses for the sequence data sets and transcriptome *de novo* assemblies were performed with CLC Genomics Workbench 4.6.1 (CLC bio, Aarhus, Denmark). Thinning was performed using 0.5 or 0.005 as the limit depending on the quality of the raw reads (visualized using FastQC; http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/), and after this thinning step, only those terminal bases with a quality score under 30 were trimmed. Global assemblies were performed on the resulting high quality reads using the default parameters of CLC Genomics Workbench 4.6.1 for *de novo* assembly.

Results

mRNA quality control

An mRNA profile is considered good in an Agilent analysis when the distribution is broad and no ribosomal RNA (rRNA) peaks are detected (Fig. 1b). However, in those

© 2011 Blackwell Publishing Ltd

Table 2 Next-generation sequencing outcome from trials of cDNA library construction. The cDNA concentration is the final concentration of fragmented cDNA libraries obtained. The profile numbers refer to those defined in Fig. S1. The Illumina outcome summarizes the number of paired-end reads (*N* reads) obtained as raw sequences and the number of assembled contigs over 2000 bp (Contigs > 2000 bp)

	Preservation method	mRNA cc (ng/μL)	A260/230	A260/280	cDNA cc (ng/μL)	Profile	*N* reads	Contigs >2000
Aphrocallistes vastus	LN_2	33.5	1.62	2.22	5.97	1	31 431 317	186*
	LN_2	43.8	2.15	1.97	1.95	4	132 339 192	388*
Cliona sp.	LN_2	28.6	1.37	2.11	10.7†	1	86 048 128	3271
	LN_2	19.4	1.28	1.65	2.62†	1	71 135 240	2821
	LN_2	18.2	1.27	2.01	0.591†	4	39 036 828	2293
Corticium candelabrum	RNA*later*	42.5	1.85	2.24	23.3	1	40 701 914	659
	LN_2	20	1.69	2.38	15.9	1	16 777 481	500
	RNA*later*	20.3	1.19	2.36	4.05†	1	88 404 620	1997
Petrosia ficiformis	LN_2	67.6	2.24	2.35	28	1	54 914 709	427
	LN_2	40.7	1.94	2.22	1.75†	4	93 040 640	1442
Spongilla lacustris	LN_2	20.6	1.48	2.07	20.4	1	192 602 980	1014*
	LN_2	74	2.23	2.22	16.3	1	40 798 603	1178
Sycon coactum	LN_2	29.5	1.88	1.99	12.5	1	103 631 158	2817
	LN_2	124.3	2.51	2.13	3.85	1	52 251 177	369
Crella elegans	RNA*later*	14.2	1.24	2.15	25	1	50 590 348	227
	RNA*later*	12.7	1.18	2.12	25.28	1	26 513 534	200

Rows coloured in grey refer to libraries run using Illumina HiSeq, while the non-coloured rows refer to cDNA libraries run with Illumina GA and GAIIx.
*The numbers correspond to contigs over 1000 bp, because in these cases the number of contigs over 2000 was unusually low because of problems with the reagents used for clustering.
†Samples prepared with the TruSeq RNA sample prep kit from Illumina.

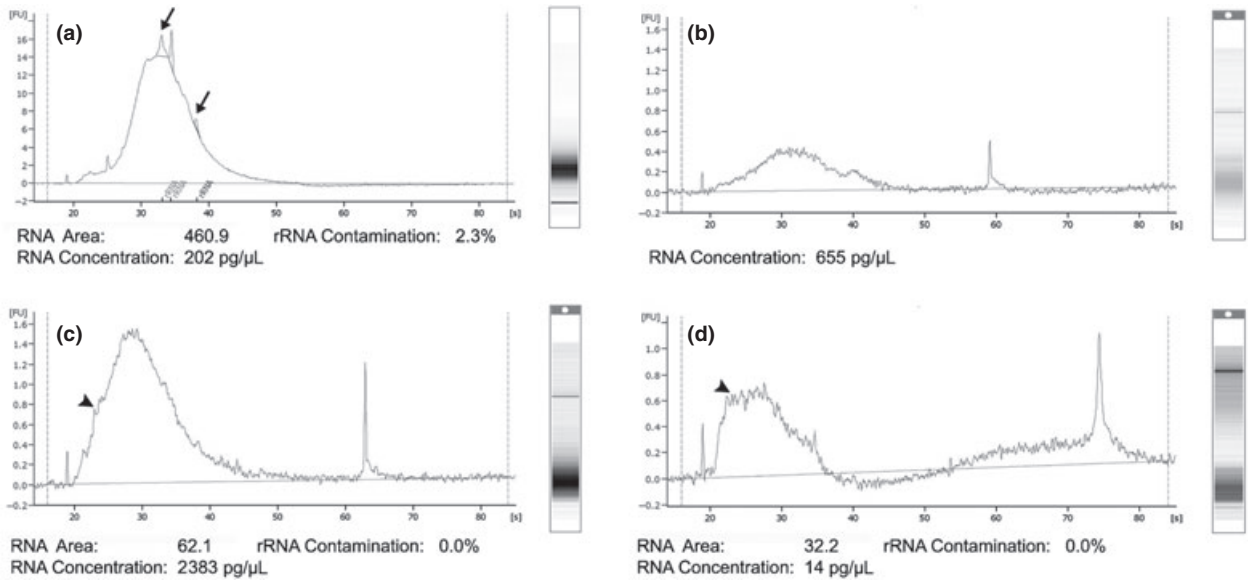

Fig. 1 Examples of quantification and quality analyses from Agilent Bioanalyzer for (a) a good quantity and quality of mRNA extracted from *Petrosia ficiformis*, (b) a poor quantity but good quality of mRNA of *P. ficiformis*, (c) a good quantity but poor quality of mRNA from *Crella elegans*, and (d) a poor quantity and quality of mRNA from *C. elegans*. Note the 18S/28S peaks (arrows) indicating rRNA contamination, and the increased number of small size fragments of mRNA (arrowheads) indicating mRNA degradation. Abbreviations: FU, fluorescence units, s, seconds.

© 2011 Blackwell Publishing Ltd

cases where rRNA contaminations of <2.5% of the total samples were assumed, cDNA libraries were sequenced successfully with very few ribosomal genes obtained (not shown).

If mRNA degradation occurs, it results in a shift in the electropherogram towards shorter fragment sizes (Fig. 1c, d). Partially degraded mRNA was only used to synthesize cDNA for *Asbestopluma occidentalis*, because we had very little tissue. However, we were not able to synthesize any cDNA out of this degraded mRNA.

On the whole, from the 85 extractions of nine sponge species (Table 1), we used the 43 mRNA extractions of best quality for cDNA synthesis for NGS, from which only 19 rendered good quality cDNA libraries (Table 2; and also Table S1, Supporting Information).

Analysing the effects of the preservation type and storage time by species

Based on the statistical analysis performed on the output of our mRNA extractions, the *t*-test indicated that the best method for RNA stabilization was the immediate flash freezing of tissues in liquid nitrogen and subsequent storage at −80 °C ($P < 0.00001$). Both quantity and quality of mRNA were higher if the tissue was flash frozen in liquid nitrogen than preserved in RNA-*later* (Fig. 2a). The analysis yielded only significant differences in the quantity and A260/230 ratio upon preservation (Fig. 2a).

When analysing the factors storage time and species separately, we only detected significant differences for

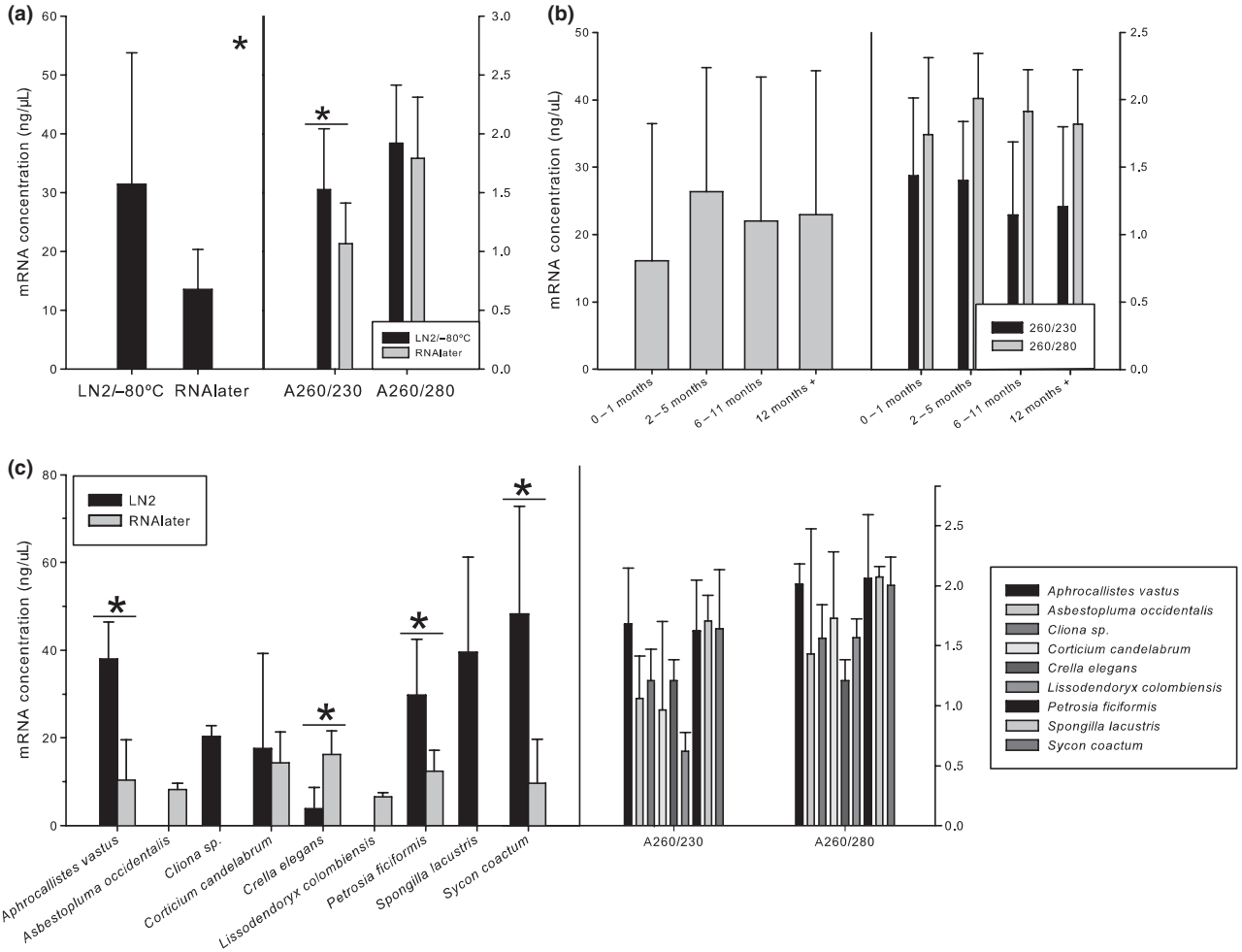

Fig. 2 Graphical representation of the statistical analysis for the effect of (a) the type of preservation on the quantity ($P < 0.05$) and quality (A260/230 ratio: $P < 0.05$; A260/280 ratio: $P > 0.05$) of mRNA; (b) the storage time periods on the quantity and quality of mRNA (none showed significant *P* values), and (c) the species on the quantity ($P < 0.05$) and quality (A260/230 ratio: $P < 0.05$; A260/280 ratio: $P > 0.05$) of mRNA. For each treatment, values represent the mean ± standard deviation. Asterisks indicate significant *P* values ($P < 0.05$). For the sake of uniformity in the graph, we depict the estimated standard deviations for *Asbestopluma occidentalis* and *Lissodendoryx colombiensis*, although only two replicates were available for each of those species.

© 2011 Blackwell Publishing Ltd

the species factor on the quantity and not the quality parameters (Fig. 2b, c).

When we analysed the effects of both the method of preservation and the storage time and the interaction between the two factors, the preservation method showed significant differences for the mRNA yield and the A260/230 ratio (Fig. 3a, b); higher values were always observed for samples preserved in liquid nitrogen. The interaction between preservation type and storage time was significant only for the A260/230 ratio (Fig. 3f); however, we performed exploratory *t*-tests to validate the assumption for all variables. Interestingly, for the quantity of mRNA, there were significant differences between both methods of preservation for periods of storage longer than 1 month. For the quality parameter A260/230, significant differences appeared for storage periods longer than 6 months. In both cases, flash-frozen samples showed higher concentrations of mRNA (Fig. 3g, h).

When performing Tukey's pairwise comparisons of the quantity and quality parameters of mRNA between the different storage times, only the A260/230 ratio showed significant differences between samples preserved in RNA*later* for 0–1 months and those preserved for longer than 12 months (MS = 0.086, d.f. = 28, $P = 0.017$) and also between tissues stored in RNA*later*

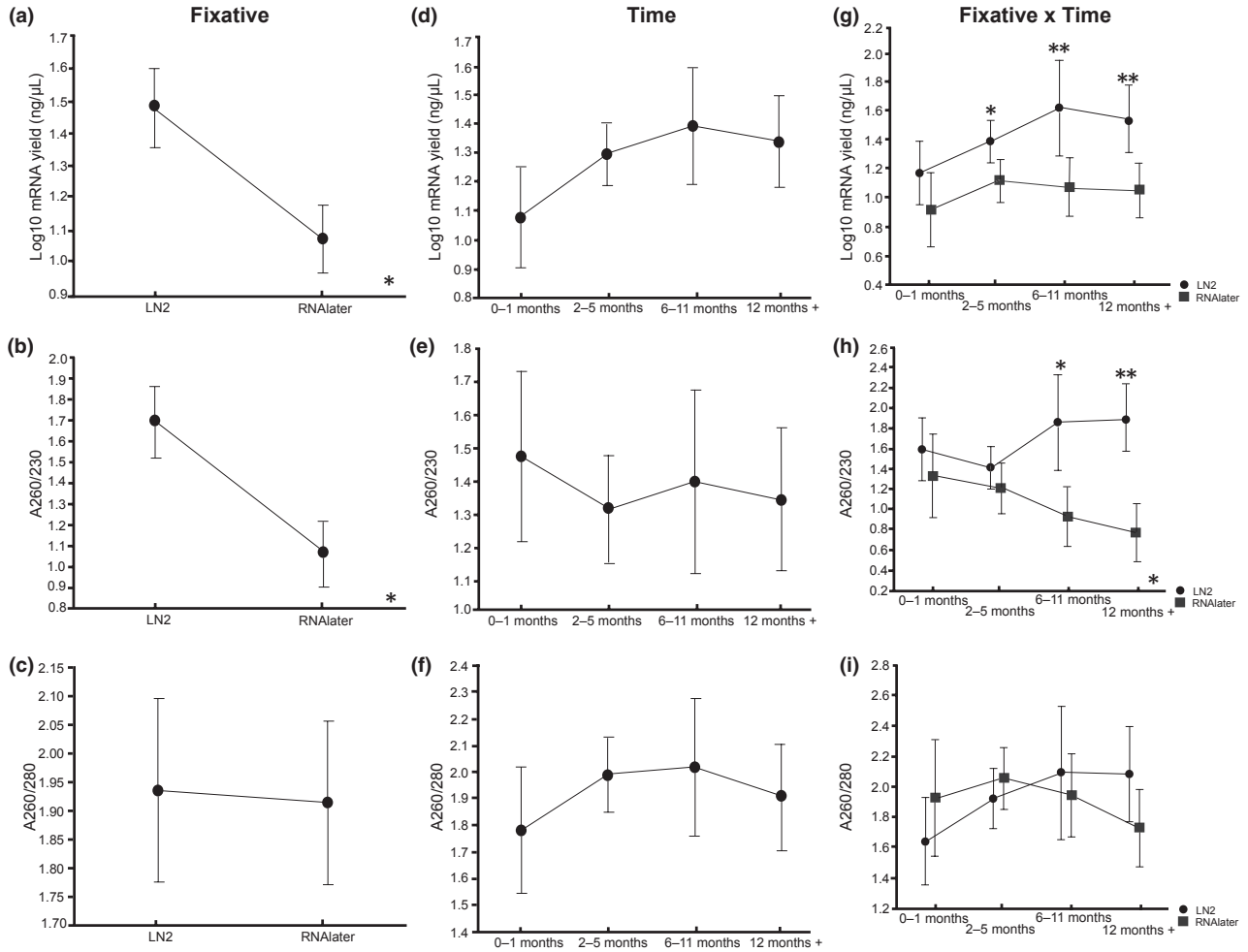

Fig. 3 Influence of preservation method (LN_2 or RNA*later*) and storage time periods (0–1, 2–5, 6–11, >12 months) on the quantity and quality of sponge mRNA. Effect of the preservation method on (a) the quantity ($P < 0.05$), (b) A260/230 ratio ($P < 0.05$) and (c) A260/280 ratio ($P < 0.05$) of mRNA. Effect of the storage time periods on (d) the quantity ($P > 0.05$), (e) A260/230 ratio ($P > 0.05$) and (f) A260/280 ratio ($P > 0.05$) of mRNA. Combined effect of the preservation method and storage time on (g) the quantity ($P > 0.05$), (h) A260/230 ratio ($P > 0.05$) and (i) A260/280 ratio ($P > 0.05$) of mRNA of mRNA. The combination of both factors was not significant ($P > 0.05$), but when individual *t*-tests were performed for each preservation type in each storage time point, differences on the quantity of mRNA between treatments were significant for 2–5, 6–11 and more than 12 months. A single asterisk over the error bars indicates a *P* value below $P < 0.05$ and double asterisk over the error bars a *P* value below $P < 0.05$ with also Bonferroni's test. Asterisks in the bottom right corner of the graph indicate significant *P* values for the ANOVAS.

© 2011 Blackwell Publishing Ltd

for 2–5 months and those stored for longer than 12 months (MS = 0.086, d.f. = 28, $P = 0.0141$).

When the factors preservation type, species studied and the interaction of both were analysed, the species factor itself rendered significant differences only for the A260/280 ratio (Fig. 4c). The preservation method, on the contrary, showed significant differences for the quantity and the quality parameter A260/230 (Fig. 4d, f). The interaction of both factors showed significant differences for all parameters (Fig. 4g–i). Exploratory *t*-tests showed that for *A. vastus*, *Sycon coactum* and *Petrosia ficiformis*, the concentration of mRNA was higher if samples were flash frozen (Fig. 4g), and for *Crella elegans* when fixed in RNA*later*. For *A. vastus* and *P. ficiformis*, the A260/230 ratio was significantly higher for flash-frozen samples (Fig. 4h), while for *C. elegans*, it was significantly higher when fixed in RNA*later*. Finally, for *A. vastus*, the A260/280 ratio was also significantly higher for flash-frozen tissues (Fig. 4i).

Tukey's *post hoc* tests revealed no significant differences in the quantity of mRNA across species. The quality parameter A260/230 showed significantly higher values for the comparison between samples fixed in RNA*later* of *A. vastus* and *C. elegans* (MS = 0.771, d.f. = 27, $P = 0.0014$). For the A260/280 ratio, there were differences between species for each preservation type. For samples preserved in LN_2, *Corticium candelabrum* showed significantly lower values than those of *A. vastus* (MS = 0.139, d.f. = 26, $P = 0.033$), *P. ficiformis* (MS = 0.139, d.f. = 26, $P = 0.044$) and *S. coactum* (MS = 0.139, d.f. = 26,

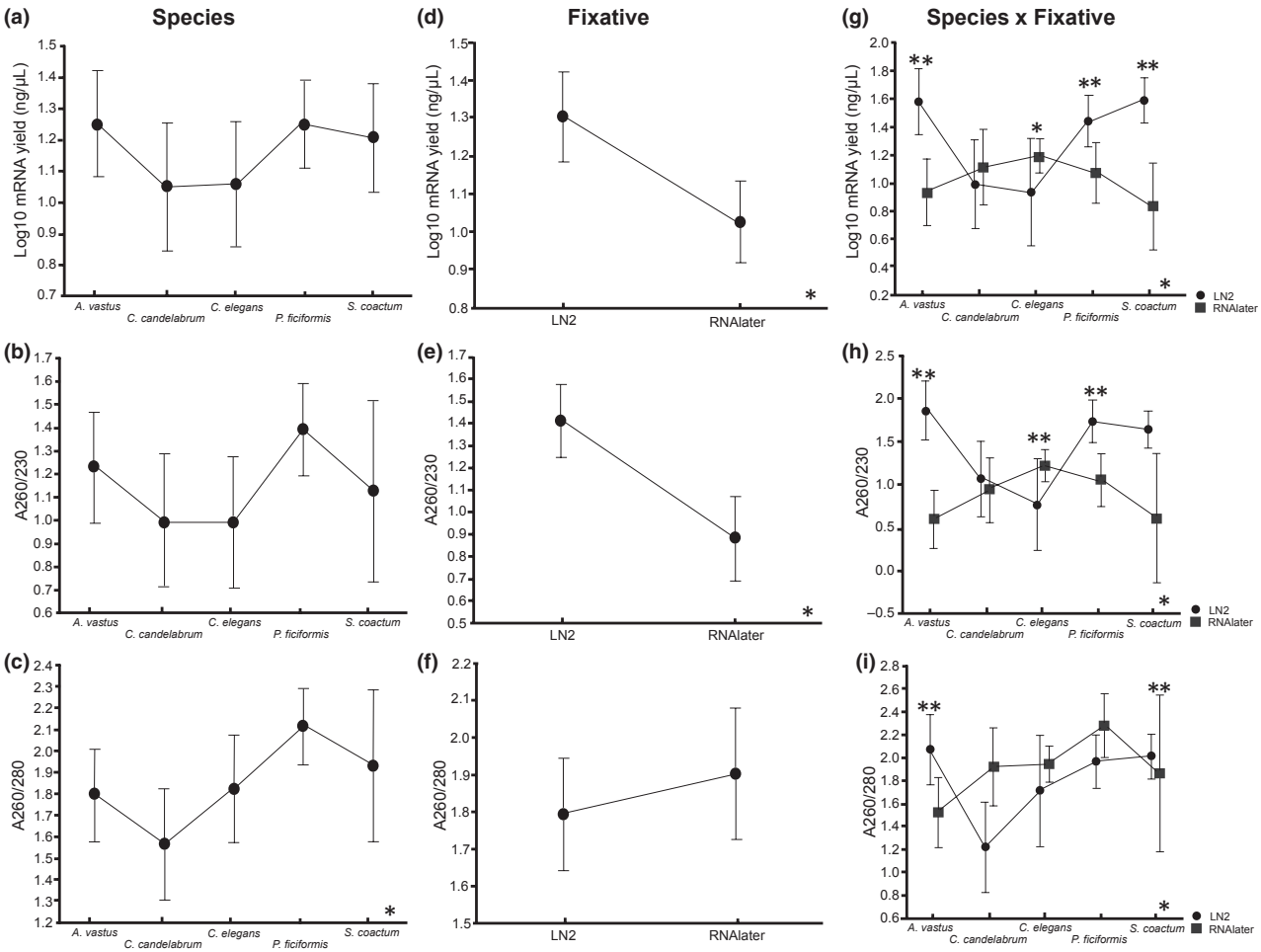

Fig. 4 Influence of preservation method (LN_2 or RN*Alater*) and species (*Aphrocallistes vastus*, *Corticium candelabrum*, *Crella elegans*, *Petrosia ficiformis* and *Sycon coactum*) on the quantity and quality of sponge mRNA. Effect of the preservation method on (a) the quantity ($P < 0.05$), (b) A260/230 ratio ($P < 0.05$) and (c) A260/280 ratio ($P > 0.05$) of mRNA. Effect of the species on (d) the quantity ($P > 0.05$), (e) A260/230 ratio ($P > 0.05$) and (f) A260/280 ratio ($P < 0.05$) of mRNA. Combined effect of the preservation method and species on (g) the quantity ($P > 0.05$), (h) A260/230 ratio ($P > 0.05$) and (i) A260/280 ratio ($P < 0.05$) of mRNA of mRNA. A single asterisk over the error bars indicates a *P* value below $P < 0.05$ when individual *t*-tests were performed, and a double asterisk over the error bars, a *P* value below $P < 0.05$, with also Bonferroni's test. Asterisks in the bottom right corner of the graph indicate significant *P* values for the ANOVAs.

© 2011 Blackwell Publishing Ltd

$P = 0.023$). For samples fixed in RNA*later*, there were only significant differences in the A260/280 of *A. vastus* and *P. ficiformis* (MS = 0.94, d.f. = 30, $P = 0.003$), RNA*later* rendering better quality results for *P. ficiformis*.

More details about the statistical results are provided in Tables S2 and S3, Supporting Information.

NGS sequencing

Nineteen cDNA libraries were successfully obtained for seven species (Table 2; and also Table S1, Supporting Information), but only 16 have been sequenced so far. mRNA yields below 12.5 ng/μL always failed when performing the first-strand synthesis (see Table S1, Supporting Information), but the limiting mRNA concentration varied between species, depending on the average quantity of mRNA obtained for each. Ideally, only cDNA libraries with a Bioanalyzer profile of type 1 (see Table S1 and Fig. S1, Supporting Information) were sent to Illumina sequencing platforms. However, in some cases, cDNA libraries with a type 4 profile were sent to Illumina HiSeq, because such platform renders 5× more reads than Illumina GA and GAIIx, and therefore, the low concentration of cDNA would be balanced out. It is important to note that the handicap in getting fragmented cDNA libraries for Illumina was always the first-strand synthesis, regardless of which preservation was used, storage time or starting mRNA concentration.

For each species, there seems to be a positive relationship between the amount of mRNA used for making the cDNA library and the number of contigs over 2000 bp obtained after assembly (Table 2). However, because building cDNA libraries is a long protocol (up to 3 days) with many different steps in which errors are carried down until the last cDNA quantitation, and libraries were synthesized by different researchers, making assumptions on the limiting mRNA quantities or qualities appears risky.

Even though the assemblies showed very different numbers of long contigs (over 2000 bp) (Table 2), all the transcriptome data are suitable for different applications such as phylogenomic analyses, probe design, gene discovery, megablast approaches or gene expression analysis.

Discussion

mRNA quality control

There is no parameter that is equivalent to the RNA integrity number (which is used to estimate total RNA quality) when analysing mRNA. The quality control is thus based on the amount of rRNA contamination and the assessment of mRNA degradation by visual estimation of increased numbers of small size fragments. This sort of interpretation can be subjective, and therefore, together with the electropherogram analysis of the extracted mRNA, the quality of the mRNA should be assessed with the A260/230 and A260/280 ratios supplied by the Nanodrop.

Good profiles of mRNA have no rRNA contamination (visualized as peaks for 18S and 28S rRNAs). However, sometimes mRNA isolations are valuable and non-repeatable, and if little rRNA contamination (i.e., <2.5%) is detected, it can be used for downstream procedures, as low recovery of ribosomal sequences is observed in the assemblies.

Maximization of A260/280 ratio of mRNA has proven crucial for subsequent experiments (Triant & Whitehead 2009). The A260/230 ratio is used less often when assessing the quality of RNA (Tattersall *et al.* 2005), but because it reflects contamination of mRNA by polyphenols, polysaccharides, buffers and other impurities that can hinder the cDNA synthesis (Becker *et al.* 2010), this parameter is important to ensure the best performance in downstream procedures. Sponges usually contain pigments, high loads of symbiotic bacteria and/or secondary compounds, and these could result in low values of A260/230. For instance, for *Crella elegans*, A260/230 ratios were always very low, regardless of RNA preservation method (maximum value 1.59). Such low values could be due to the high amount of mucus contained in the tissue, and therefore, optimization of A260/230 ratios for this species was not possible.

Optimal preservation and storage time for sponge tissue samples

Preservation in liquid nitrogen followed by storage at −80 °C rendered the best results in terms of quantity (nanogram per microlitre) and A260/230 ratio of mRNA for every species. Samples from the same species preserved in liquid nitrogen sometimes had twice the amount of mRNA of those stored in RNA*later* (Fig. 2c). Liquid nitrogen flash freezing and storage at −80 °C has proven ideal for preserving microcrustaceans (Gorokhova 2005), sponges (Simister *et al.* 2011) and human tissues (Wang *et al.* 2006), but RNA*later* seems to be as efficient as flash freezing regarding the quality of mRNA, as also observed recently for a wide variety of other metazoans (Mutter *et al.* 2004; Gayral *et al.* 2011). Even though liquid nitrogen seems the optimal preservation method for obtaining large quantities of mRNA, unfortunately obtaining and keeping liquid nitrogen in remote locations is not easy and −80 °C freezers are uncommon on ships or even at small marine stations where invertebrate marine samples are often collected and prepared. Therefore, even though the RNA*later* stabilization of RNA in sponges renders less mRNA, the mRNA yield is

© 2011 Blackwell Publishing Ltd

usually enough to construct cDNA libraries for NGS when manufacturer's instructions are followed carefully, and if samples are processed within a month. Also, RNA*later* allows more flexibility in the storage and transport because samples can be placed at room temperature for a few days. It is important to note that even though for the poecilosclerid *C. elegans* RNA*later* seemed to perform better than flash frozen samples, for other poecilosclerids this may not be the case, as suggested by the poor yields of mRNA obtained from RNA*later*-fixed samples from the poecilosclerids *Asbestopluma occidentalis* and *Lissodendoryx colombiensis*.

Long storage of tissue samples could potentially be an issue when preparing cDNA for sequencing purposes. However, our statistical analyses show that there are no significant differences between storing samples for up to 1 month or more than 12 months. However, when the effect of the different preservation methods is considered, the quantity of mRNA recovered from samples flash frozen in liquid nitrogen was significantly higher than for tissues stabilized in RNA*later* for periods longer than 2 months up to 12 months. If samples are stored for <1 month, both preservation methods performed equally efficiently, even if RNA*later* stored samples were not stored at −80 °C for short periods of time (up to a few days). Therefore, if tissues are to be processed in <1 month, either preservation method works equally well for mRNA stabilization, but flash freezing in liquid nitrogen should be preferred if samples are to be processed after longer periods of time.

The amount of mRNA recovered depended on the sponge species, which could be due to different cell/spicule ratios for the different taxa. But when the effect of the preservation was taken into account, we detected no significant differences in the yield and A260/230 ratio of the mRNA extracted from five different sponges, and only the A260/280 was significantly different as a function of the species. To be consistent, the amount of tissue preserved, and therefore the amount of mRNA recovered, should take into account the relative amount of spicules and cell contents per sample. For instance, in the case of *Aphrocallistes vastus, Sycon coactum* or *Petrosia ficiformis*, more than 200 mg were fixed in each tube, in an attempt to obtain the targeted 80 mg of tissue and to recover the same amount of mRNA as in less densely spiculated sponges, given that from 50% to 80% of the weight was because of spicule content.

The influence of the extraction method on the quantity and quality of mRNA was not tested statistically because of differences in sample sizes, essentially because some extraction methods rendered minuscule amounts of mRNA and were quickly discarded. Therefore, although the Dynabeads® mRNA DIRECT™ Kit seemed to perform more efficiently and more consistently for our sponge samples, it is important to note that for *Corticium candelabrum*, a standard trizol protocol (using Tri-Reagent) followed by a purification of mRNA with the Dynabeads® mRNA Purification Kit for mRNA Purification from Total RNA preps always produced better mRNA yields and quality than Dynabeads® mRNA DIRECT™ Kit. As noted before, sponges usually contain pigments, many also have high loads of symbiotic bacteria and/or secondary compounds that could hinder mRNA extractions, and these might interfere with the extraction reagents and yields/quality obtained for certain species. Therefore, we encourage the assessment of different extraction methods to find the ideal for each tissue type. For instance, Simister *et al.* (2011) found that the extraction of total RNA using the method developed by Griffiths *et al.* (2000) or a trizol-based extraction performed highly efficiently for sponge tissues.

Here, we show that there are no significant differences in the mRNA quality (A260/280 ratio) with the preservation mode (flash frozen vs. RNA*later*) but that there are differences in terms of the quantity and the A260/230 ratio. Also, we found that if tissue was stored for periods longer than 1 month, then preservation by flash freezing in liquid nitrogen gave a better mRNA yield than for those tissues stabilized in RNA*later*. This may be a peculiarity of sponges and the reactions that spicules may have with the reagents in RNA*later*, so further study of non-sponge organisms could add information to this issue. However, both preservation methods, easily available to all researchers, seem to be good for NGS study of sponges.

Acknowledgements

We thank Andi Hejnol, Pamela J. Windsor, Sónia Andrade, and Freya Goetz for help with mRNA extraction; Sònia de Caralt, Iosune Uriz, Andrea Blanquer, Janina González, Sergio Taboada, and Joan Mora for help in field sampling, and Vanessa González for critical reading and valuable suggestions to the manuscript. We are also thankful to April and Malcolm Hill for supplying *Cliona* sp. tissue, and Steve Vollmer for providing the protocol for building cDNA libraries. This work has been supported by a Marie Curie Fellowship to A.R., a FPI pre-doctoral Fellowship (BES-2008-003009) of the Spanish Ministry of Science and Innovation and a grant included in the project Benthomics (CTM2010-22218-C02-01) to A.R.R.-P., and a Fulbright of the Spanish Ministry of Science and Innovation (EX2009-0128) to S.C. This material is based upon work supported by a grant of the Marie Curie Outgoing Fellowship Program to A.R., by the National Science Foundation under Grant No. DEB #0844881 to G.G, and NSERC Discovery and Ship Time grants to S.P.L.

References

Adamska M, Degnan SM, Green KM *et al.* (2007) Wnt and TGF-ß expression in the sponge *Amphimedon queenslandica* and the origin of metazoan embryonic patterning. *PLoS ONE*, **2**, e1031.

© 2011 Blackwell Publishing Ltd

Adamska M, Larroux C, Adamski M *et al.* (2010) Structure and expression of conserved Wnt pathway components in the demosponge *Amphimedon queenslandica*. *Evolution & Development*, **12**, 494–518.
Becker C, Hammerle-Fickinger A, Riedmaier I, Pfaffl MW (2010) mRNA and microRNA quality control for RT-qPCR analysis. *Methods*, **50**, 237–243.
Blow N (2009) Transcriptomics: the digital generation. *Nature*, **458**, 239–242.
Dunn CW, Hejnol A, Matus DQ *et al.* (2008) Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature*, **452**, 745–749.
Ekblom R, Galindo J (2011) Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*, **107**, 1–15. doi: 10.1038/hdy.2010.152.
Elmer KR, Fan S, Gunter HM *et al.* (2010) Rapid evolution and selection inferred from the transcriptomes of sympatric crater lake cichlid fishes. *Molecular Ecology*, **19**, 197–211.
Funayama N, Nakatsukasa M, Mohri K, Masuda Y, Agata K (2010) Piwi expression in archeocytes and choanocytes in demosponges: insights into the stem cell system in demosponges. *Evolution & Development*, **12**, 275–287.
Gayral P, Weinert L, Chiari Y *et al.* (2011) Next-generation sequencing of transcriptomes: a guide to RNA isolation in nonmodel animals. *Molecular Ecology Resources*, **11**, 650–661. doi: 10.1111/j.1755-0998.2011.03010.x
Gilad Y, Pritchard JK, Thornton K (2009) Characterizing natural variation using next-generation sequencing technologies. *Trends in Genetics*, **25**, 463–471.
Gorokhova E (2005) Effects of preservation and storage of microcrustaceans in RNA*later* on RNA and DNA degradation. *Limnology and Oceanography, Methods*, **3**, 143–148.
Griffiths RI, Whiteley AS, O'Donnell AG, Bailey MJ (2000) Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA and rRNA-based microbial community composition. *Applied and Environmental Microbiology*, **66**, 5488–5491.
Harcet M, Roller M, Cetkovic H *et al.* (2010) Demosponge EST sequencing reveals a complex genetic toolkit of the simplest metazoans. *Molecular Biology and Evolution*, **27**, 2747–2756.
Hejnol A, Obst M, Stamatakis A *et al.* (2009) Assessing the root of bilaterian animals with scalable phylogenomic methods. *Proceedings of the Royal Society. B, Biological Sciences*, **276**, 4261–4270.
Larroux C, Fahey B, Degnan SM *et al.* (2007) The NK homeobox gene cluster predates the origin of Hox genes. *Current Biology*, **17**, 706–710.
Leys SP, Riesgo A (in press) Epithelia, an evolutionary novelty of metazoans. *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution*, doi: 10.1002/jez.b.21442, in press.
Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y (2008) RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Research*, **18**, 1509–1517.
Morozova O, Marra MA (2008) Applications of next-generation sequencing technologies in functional genomics. *Genomics*, **92**, 255–264.
Mutter GL, Zahrieh D, Liu C *et al.* (2004) Comparison of frozen and RNA-Later solid tissue storage methods for use in RNA expression microarrays. *BMC Genomics*, **5**, 88.
Nichols SA, Dirks W, Pearse JS, King N (2006) Early evolution of animal cell signaling and adhesion genes. *Proceedings of the National Academy of Sciences*, **103**, 12451–12456.
Philippe H, Derelle R, Lopez P *et al.* (2009) Phylogenomics revives traditional views on deep animal relationships. *Current Biology*, **19**, 706–712.
Reis-Filho JS (2009) Next-generation sequencing. *Breast Cancer Research*, **11**, 1–7.
Rivera AS, Hammel JU, Haen KM *et al.* (2011) RNA interference in marine and freshwater sponges: acting knockdown in *Tethya wilhelma* and *Ephydatia muelleri* by ingested dsRNA expressing bacteria. *BMC Biotechnology*, **11**, 67–82.
Shendure J, Ji H (2008) Next-generation DNA sequencing. *Nature Biotechnology*, **26**, 1135–1145.
Simister RL, Schmitt S, Taylor MW (2011) Evaluating methods for the preservation and extraction of DNA and RNA for analysis of microbial communities in marine sponges. *Journal of Experimental Marine Biology and Ecology*, **397**, 38–43.
Simpson JT, Wong K, Jackman SD *et al.* (2009) ABySS: a parallel assembler for short read sequence data. *Genome Research*, **19**, 1117–1123.
Srivastava M, Simakov O, Chapman J *et al.* (2010) The *Amphimedon queenslandica* genome and the evolution of animal complexity. *Nature*, **466**, 720–726.
Sullivan J, Sher D, Eisenstein M *et al.* (2008) The evolutionary origin of the Runx/CBFbeta transcription factors – studies of the most basal metazoans. *BMC Evolutionary Biology*, **8**, 228.
Sultan M, Schulz MH, Richard H *et al.* (2008) A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science*, **321**, 956–960.
Tattersall EAR, Ergul A, AlKayal F *et al.* (2005) Comparison of methods for isolating high-quality RNA from leaves of grapevine. *American Journal of Enology and Viticulture*, **56**, 400–406.
Toth AL, Varala K, Newman TC *et al.* (2007) Wasp gene expression supports an evolutionary link between maternal behavior and eusociality. *Science*, **318**, 441–444.
Triant DA, Whitehead A (2009) Simultaneous extraction of high-quality RNA and DNA from small tissue samples. *Journal of Heredity*, **100**, 246–250.
Vera JC, Wheat CW, Fescemyer HW *et al.* (2008) Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. *Molecular Ecology*, **17**, 1636–1647.
Wang SS, Sherman ME, Rader JS, Carreon J, Schiffman MM, Baker CC (2006) Cervical tissue collection methods for RNA preservation: comparison of snap-frozen, ethanol-fixed, and RNAlater-fixation. *Diagnostic Molecular Pathology*, **15**, 144–148.
Zerbino DR, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*, **18**, 821–829.

Data accessibility

The information regarding concentration and quality of mRNA of all the extractions performed during the study is available as Supporting Information.

Supporting Information

Additional supporting information may be found in the online version of this article.

Data S1 Material and methods.

Fig. S1 Bioanalyzer profiles for double stranded cDNA libraries for sponge samples showing (a) a tight band of targeted size with high cDNA concentration (profile 1), (b) a tight band of targeted size with 'bumps' of smaller (in this case) or larger fragments; i.e., poor quality (profile 2), (c) no bands (profile 3), (d) a tight band of targeted size with low cDNA concentration (profile 4).

Table S1 Next Generation Sequencing outcome from trials of cDNA library construction.

Table S2 Statistical results for all the analyses performed to assess the effects of preservation method (fixative), storage time, and species on the quantity and quality of mRNA.

Table S3 Means ± standard deviations of the mRNA yield and the quality parameters A260/230 and A260/280 ratios for all variables studied.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

© 2011 Blackwell Publishing Ltd