Low-cost RNA extraction method for highly scalable transcriptome studies

Kanami Yoshino, Ryo Nishijima and Taiji Kawakatsu*

Institute of Agrobiological Sciences, National Agriculture and Food Research Organization, 3-1-3 Kan-nondai, Tsukuba, Ibaraki 305-8604, Japan

RNA extraction has been improved by integration of a variety of materials in the protocol, such as phenol, guanidine thiocyanate, and silica, according to the case-specific demands. However, few methods have been designed for high-throughput RNA preparation for large-scale transcriptome studies. In this study, we established a high-throughput guanidinium thiocyanate and isopropyl alcohol based RNA extraction method (HighGI). HighGI is based on simple and phenol-free homemade buffers and the cost is substantially lower than a column-based commercial kit. We demonstrated that the quality and quantity of RNA extracted with HighGI were comparable to those extracted with a conventional phenol/chloroform-based method and a column-based commercial kit. HighGI retained small RNAs less than 200 bp, which are lost with a commercial column-based kit. We also demonstrated that HighGI is readily applicable to semi-automated RNA extraction. HighGI enables high-throughput RNA extraction for large-scale RNA preparation with high yield and quality.

Key Words: high-throughput RNA extraction, guanidinium thiocyanate, magnetic beads, small RNA extraction.

Introduction

Gene expression represents an essential biological activity and reflects specific biological reactions in cells. For gene expression analysis, RNA extraction is a crucial procedure for a variety of downstream analyses, such as RNA sequencing. The decline in cost of high-throughput sequencing has enabled large-scale transcriptome studies to reveal novel gene functions and interactions among gene regulatory networks (Yang et al. 2019). Progress in high-throughput sequencing technologies has been so rapid that sample preparation is becoming a bottleneck with regard to cost and labor. Recently, several high-throughput and cost-effective RNA-sequencing library preparation methods have been reported (Kamitani et al. 2019, Townsley et al. 2015). A number of RNA extraction methods, ranging from classic methods to improved procedures, have been developed. Phenol/chloroform extraction and acid guanidine thiocyanate–phenol–chloroform extraction are conventional RNA extraction methods (Chomczynski and Sacchi 1987, Logemann et al. 1987, MacDonald et al. 1977). Numerous commercial kits involving phenol- and guanidine-based protocols are available (Chomczynski et al. 2010). Although RNA extracted using these methods are of high quality and have been widely used for diverse studies, phenol should be used with particular care because of its corrosive, toxic, and health hazardous nature. An extraction procedure based on phenol-free cetyltrimethylammonium bromide and polyvinylpyrrolidone has also been adopted, especially for tissues that contain high levels of secondary metabolites, including polysaccharides and polyphenols (Barbier et al. 2019, Chang et al. 1993, Jaakola et al. 2001, Zeng and Yang 2002). Although cetyltrimethylammonium bromide and polyvinylpyrrolidone-based extraction is useful for highly viscous materials, ethanol or isopropyl alcohol (IPA) precipitation with prolonged centrifugation steps is required. Combination of guanidinium thiocyanate (GTC) and reducing agents, such as dithiothreitol and 2-mercaptoethanol, is among the most commonly used RNA extraction methods (Chirgwin et al. 1979, Ullrich et al. 1977). Given that GTC is a chaotropic salt, the high concentration of GTC facilitates binding of the nucleic acids to a silica membrane column or silica-coated magnetic beads. Therefore, commercial kits that use GTC and a silica membrane column or silica-coated magnetic beads eliminate the requirement for alcohol precipitation steps and thus enable rapid RNA extraction. However, the use of commercial kits is expensive and column-based kits are unsuitable for processing a large number of samples. In this study, we developed a simple and low-cost RNA extraction method, termed high-throughput GTC and IPA based RNA extraction (HighGI),
and compared the performance of the procedure with those of a conventional phenol/chloroform-based method and a column-based commercial kit, assessed with regard to yields, purity, integrity and composition of RNA. The yield and quality of RNA extracted by HighGI were not inferior to those attained using a conventional extraction method or a commercial kit. Furthermore, the HighGI protocol also isolated small RNAs of less than 200 bp, which are lost with a commercial column-based kit. We demonstrated that HighGI is readily applicable for automated RNA extraction using a liquid handling system for large-scale transcriptome studies.

Materials and Methods

Plant materials

We collected rice (Oryza sativa L.) shoots at 6 days after germination from seedlings grown on Murashige and Skoog medium, and root tips at 8 weeks after germination from plants grown in a growth chamber. We also collected leaves at 5 months after germination from plants grown in a paddy field. The fresh weights of the harvested shoots and leaves were approximately 40 and 30 mg, respectively, and five individual root tips were used. Samples were flash frozen in liquid nitrogen, ground to a fine powder with a multi-beads shoker (Yasui-Kikai), and stored at −80°C until RNA extraction.

HighGI RNA extraction

The experimental workflow for the HighGI RNA extraction method is shown in Fig. 1. Frozen samples were suspended in 200–300 μL lysis buffer [4 M GTC, 1 mM Tris-HCl, 0.05% Tween 20, pH 7.5]. The entire following process was conducted at room temperature. The lysate was centrifuged at 12,000 × g for 2 min and 150 μL of the supernatant was transferred to a new 1.5 mL tube or a 2 mL 96-deep-well plate. An aliquot (270 μL) of the solid-phase reversible immobilization beads suspension [SeraMag Speedbeads in 18% PEG8000, 1 M NaCl, 20 mM Tris-HCl, 10 mM EDTA, 0.05% Tween 20, pH 8 (Rohland and Reich 2012)] and 270 μL IPA were added to the supernatant and left for 5 min for RNA to bind to the magnetic beads. Commercial solid-phase reversible immobilization beads, such as AMPure XP beads, would work as well. The magnetic beads were captured with a magnetic stand or plate and the supernatant was discarded. The magnetic beads were washed with 300 μL wash buffer [1 M GTC, 1 mM Tris-HCl, 0.05% Tween 20, 1 M NaCl, 40% IPA, pH 7.5] or 85% ethanol. The magnetic beads were washed three times with 690–900 μL of 85% ethanol. The magnetic beads were air-dried for 5 min on the magnetic stand and RNA was eluted in 100 μL nuclease-free water. For automated RNA extraction, the procedures after lysate preparation were automated on a Biomek i7 workstation (Beckman Coulter).

Phenol/chloroform-based RNA extraction

Frozen ground samples were homogenized in 400 μL RNA extraction buffer [50 mM Tris-HCl (pH 9.0), 1% SDS, 0.1 M NaCl, 5 mM EDTA-2Na]. The samples were deproteinized with 400 μL phenol/chloroform (1:1, v/v) at least twice, then washed with 400 μL chloroform. Nucleic acids were precipitated with 0.1 volume of 3 M sodium acetate (pH 5.2) and 1.0 volume of IPA. The samples were flash frozen with liquid nitrogen, then centrifuged for efficient small RNA precipitation.

RNA extraction with a commercial kit

Total RNA was extracted using the RNeasy® Plant Mini Kit (Qiagen) following the manufacturer’s protocol (Fig. 1).

Quantification of RNA yield and integrity

The concentration and purity of the products were assessed by measuring the absorbance at 260 nm and absorbance ratios (260/280 and 260/230 nm) using a NanoDrop™ 2000 spectrophotometer (Thermo Scientific). The integrity of total RNA was analyzed with an Agilent 2100 Bioanalyzer System (Agilent Technologies) using the Plant RNA Nano Kit by measuring the molecular size distribution and RNA integrity number (RIN), ranging from 1 to 10. A low RIN indicates degradation of the RNA molecules. Given that the amount of collected supernatants after the first centrifugation (150 μL out of 200–300 μL for HighGI, all ~400 μL for the phenol/chloroform method, and all ~450 μL for RNeasy kit extraction), we estimated possible recovery yields by multiplying the reciprocal of the collection ratio (i.e. total volume/colllected volume of...
the supernatant). Statistical analysis of the yield of nucleic acids was performed using a multiple comparison test (Tukey–Kramer method, \( p < 0.01 \)).

**RT-qPCR**

Total RNAs were extracted from the same shoot samples with Phenol-SDS, RNeasy and HighGI methods, with three biological replicates (i.e. three samples \( \times \) three methods). Total RNAs (400 ng) extracted from shoots were used for cDNA synthesis using ReverTra Ace qPCR RT Master Mix with a gDNA Remover Kit (Toyobo) following to the manufacturer’s protocol. qPCR reactions were conducted in the Viia 7 Real-Time PCR System (Applied Biosystems) using SYBR Premix Ex Taq II kit (Takara) with 20 μL reaction volume, with the following conditions: 95°C for 30 s, followed by 40 cycles of 95°C for 5 s, 60°C for 30 s and 72°C for 30 s, with three biological replicates. Primers used were listed in *Table 2*.

**Results and Discussion**

**Optimization of HighGI RNA extraction**

Critical steps for RNA extraction using magnetic beads are 1) preparation of the lysate containing intact RNAs, 2) binding of RNAs to the magnetic beads, and 3) washing out impurities. We used shoots as input samples for following optimization. First, we optimized the GTC concentration in the lysis buffer. The acid guanidine thiocyanate–phenol–chloroform method uses 4 M GTC as the deproteinization agent. Given that preparation of GTC at a concentration higher than 4 M is possible but difficult, we compared the concentration of GTC in the lysis buffer (2 M or 4 M). The yield with 2 M GTC was half of that with 4 M GTC (22.2–35.6 μg vs 80.9 ± 8.3 μg; \( n = 2 \) and 3, respectively). This result suggested that the higher concentration of GTC more efficiently deproteinized and/or enhanced the binding of RNA to silica-coated magnetic beads. Therefore, we decided to use 4 M GTC for the lysis buffer.

Second, we optimized the binding buffer containing magnetic beads. We used solid-phase reversible immobilization beads, which are used for nucleic acid purification during high-throughput sequencing of RNA libraries, for two reasons: the simplicity and the presence of PEG8000, which can enhance nucleic acid binding to silica-coated magnetic beads. To maximize the binding efficiency, we included IPA in the wash buffer.

Third, we considered the wash buffer. The GTC-based commercial kits, such as RNeasy and the Agencourt Chloropure Kit (Beckman), include GTC in the wash buffer, but the concentration used is confidential. Given that GTC strongly denatures proteins, excessive carryover of GTC should be avoided. In addition, GTC carryover decreases the 260/230 absorbance ratio. In the comparison of GTC concentrations (0.5–2.0 M) without NaCl and with 2 M GTC for lysis buffer, the highest yields were obtained with 1 M GTC for the wash buffer (9.5 ± 2.7 for 2.0 M GTC, 13.2 ± 2.1 for 1.6 M GTC, 23.0 ± 6.3 for 1 M GTC, and 18.7 ± 4.6 for 0.5 M GTC). In addition, we compared 1 M GTC and 85% ethanol for the wash buffer. Eighty-five percent ethanol yielded the higher amount of RNA with the comparable absorbance ratio compared with that using 1 M GTC. The RIN with 85% ethanol was marginally lower than that with 1 M GTC. Therefore, we selected 1 M GTC as the primary wash buffer, and 85% ethanol as the optional wash buffer.

**Comparison with conventional RNA extraction**

**Quantity and quality**

We extracted RNAs from young shoots, mature leaves, and root tips using three methods: phenol-SDS, RNeasy, and HighGI. From all tissues, the phenol-SDS method yielded the highest amounts of RNA (*Fig. 2*). The HighGI method yielded the lowest amounts of RNA. However, the estimated possible recovery amounts, when the entire supernatant was used for RNA extraction, was higher than those attained with RNeasy (*Fig. 2*). The afore-mentioned experiments omitted DNase I treatment, therefore some genomic DNA was included in the extract, which may be a problem for some downstream analyses. We prepared the optional DNase I treatment step after the first wash step. DNase I treatment decreased the amounts of RNA, likely due to elimination of contaminated genomic DNA, but the quality was maintained (*Table 1, Fig. 3*).

**Small RNA extraction**

The RIN was originally formulated for humans and mouse, and largely depends on the peak pattern around 17S and 28S ribosomal RNAs (rRNAs). Plant shoots and leaves contain abundant chloroplasts. Chloroplast rRNA peaks are located around those of 17S and 28S rRNAs, leading to a lower RIN in shoots and leaves. In addition, peaks detected
in smaller regions decrease the RIN because they are recognized as degraded RNAs. We observed substantial amounts of peaks in RNAs smaller than 200 nucleotides (nt) extracted with phenol-SDS and HighGI, although rRNA degradation was not observed (Fig. 3). Given that RNAs extracted with the phenol-SDS method contain small RNAs, including micro RNAs and short interfering RNA (Kawakatsu et al. 2012), these peaks likely represent small RNAs. This result shows that RNAs extracted with HighGI also contain small RNAs. Therefore, RNA extracted with HighGI can be used for small RNA sequencing, which is an important analytic procedure in epigenetics. RNAs extracted with RNeasy showed the 5S rRNA peak, but no peak was observed in smaller regions. Both HighGI and RNeasy use GTC and silica, but only HighGI can collect small RNAs. This outcome can be explained by the different alcohol in the buffers: IPA in HighGI and ethanol in RNeasy. Given that IPA shows lower polarity, the binding of hydrophilic nucleic acids to silica was likely promoted.

**Composition of mRNAs**

To compare the compositions of mRNAs in RNAs extracted from shoots with Phenol-SDS, RNeasy and HighGI methods, we quantified the abundance of randomly selected six transcripts with different lengths (0.5–6.2 kb; Table 2) by RT-qPCR. We designed the primers on exons as close to transcription start sites as possible to confirm, at least nearly, full-length transcripts were captured (Supplemental Fig. 1). The abundance of all examined transcripts was not significantly differed among samples (paired Student’s t-Test, p > 0.05), except for gene 5 (Fig. 4). However, the fold change of gene 5 was at most 1.6, which was lower than the criteria commonly used for differential expression analysis based on RNA-seq or microarray. Therefore, we concluded the compositions of mRNAs were comparable (Fig. 4). Although RNAs extracted with Phenol-SDS and HighGI methods contained RNAs smaller than 200 nt, these RNAs (<200 nt) were removed during mRNA-seq library preparation procedures. Indeed, mRNA-seq data from RNAs extracted with RNeasy (only >200 nt transcripts) and with

---

**Table 1.** Comparison of absorbance (nm) ratio and RNA integrity number (RIN)

|               | Shoots | Leaves | Root tips |
|---------------|--------|--------|-----------|
|               | 260/280| 260/230| RIN       | 260/280| 260/230| RIN       | 260/280| 260/230| RIN       |
| Phenol & SDS  | 2.0 ± 0.0 | 2.5 ± 0.0 | 8.0 | 1.9 ± 0.0 | 2.4 ± 0.0 | 7.0 | 1.9 ± 0.0 | 2.5 ± 0.0 | 9.2 |
| RNeasy Plant Mini Kit | 2.1 ± 0.0 | 2.4 ± 0.0 | 8.1 | 2.1 ± 0.0 | 2.3 ± 0.1 | 7.3 | 2.1 ± 0.0 | 2.2 ± 0.1 | 8.8 |
| HighGI       | 1.9 ± 0.0 | 1.8 ± 0.3 | 7.8 | 1.9 ± 0.0 | 1.7 ± 0.3 | 7.6 | 2.0 ± 0.0 | 1.9 ± 0.4 | 9.6 |
| HighGI (85% ethanol) | 1.9 ± 0.0 | 1.8 ± 0.2 | 7.7 | 2.0 ± 0.0 | 1.8 ± 0.1 | 7.2 | 1.9 ± 0.1 | 1.9 ± 0.6 | 9.1 |
| HighGI (+DNase) | 2.0 ± 0.0 | 2.0 ± 0.0 | 8.4 | 2.0 ± 0.0 | 1.7 ± 0.3 | 7.6 | n.a.       | n.a.       | n.a.       |
| RNeasy buffer (Biomek) | 2.0 ± 0.0 | 2.3 ± 0.1 | 8.3 | n.a.       | n.a.       | n.a.       | n.a.       | n.a.       | n.a.       |
| HighGI (Biomek) | 2.0 ± 0.0 | 2.1 ± 0.1 | 8.2 | n.a.       | n.a.       | n.a.       | n.a.       | n.a.       | n.a.       |

*a* 85% ethanol was used as the wash buffer. Absorbance ratio is the average of replicate measurements (n = 8 for application of Biomek, n = 3 for the other methods, ± standard deviation). The RIN value was measured for one sample.

n.a.: not applicable.

---

**Fig. 3.** Electropherograms of total RNA extracted using each method. The y-axis represents fluorescence intensity and the x-axis represents the length of the RNA molecule (nucleotides; nt). The RNA integrity number (RIN) values are indicated. The lowest peak at 25 nt is a size marker.
Trizol (including <200 nt transcripts) were highly correlated (Pearson correlation coefficient >0.99; Sultan et al. 2014). Therefore, we anticipate that RNA extraction methods would have little impact on mRNA-seq data.

Feasibility for large-scale studies

Initially, we developed HighGI using 1.5 mL tubes. To achieve high-throughput RNA extraction, we performed HighGI on 96-well 2 mL deep-well plates using a liquid handling system. We also performed the modified HighGI, in which lysis buffer and wash buffer were substituted with RLT buffer and RW1 buffer from the RNeasy kit, respectively. The amounts and qualities of RNAs extracted with HighGI and modified HighGI on 96-well deep-well plates were comparable to those of RNAs extracted with HighGI in 1.5 mL tubes (Table 1, Fig. 2). We utilized the liquid handling system for HighGI on 96-well plates, but one can perform the procedure with multi-channel pipettes.

Conclusion

In this study we developed the simple HighGI RNA extraction method. We demonstrated that the quantity and quality of RNA extracted with HighGI were comparable to conventional methods or commercial kits. We also demonstrated that HighGI can be used for small RNA extraction. In addition, HighGI can be semi-automated using liquid handling systems to process a large number of samples and high-throughput sequencing. The cost of large-scale studies may be a bottleneck. We estimated that the cost of HighGI is much cheaper than commercial kits, at JPY 31 per sample, which is approximately 1/27 the cost of the RNeasy Plant Mini Kit and 1/4 the cost of the Agencourt Chloropure Kit (Supplemental Tables 1, 2). We have tested only rice shoots, leaves, and root tips, but we anticipate that HighGI can be applied to a variety of sample types, for which GTC-based commercial kits are used for RNA extraction. Given that GTC gelatinizes starch, HighGI cannot be applied for starch-rich samples, such as endosperm. HighGI can be used for routine RNA extraction and will be useful for RNA extraction for highly scalable transcriptome studies.

Author Contribution Statement

KY performed the experiments and wrote the manuscript; RN performed the experiments; TK conceived and designed the research, and wrote the manuscript.

| Gene no. | Locus Identifier | Transcript length (bp) | Forward and reverse primer sequence (5ʹ-3ʹ) |
|---------|-----------------|------------------------|-------------------------------------------|
| 1       | RUBQ2 (LOC_Os02g06640) | 1,451 | F: GAGCCTCTGTTGTCGTAAGTA R: ACTCGATGCTCCATTAACCC |
| 2       | LOC_Os07g23990 | 6,206 | F: GCCGCTATGCGTAAACCCGG R: ATCCGGACACATCTCCTCAAGC |
| 3       | LOC_Os07g46460 | 5,804 | F: AGGTTCTGGAATGGCTTTTCR: AGGCTTCTGTCAGTGAATACC |
| 4       | LOC_Os06g40940 | 3,624 | F: ACCAGCAACATCTGCACCTGCC R: AGTCCATGTACCGGTGCGCAAA |
| 5       | LOC_Os03g52840 | 2,283 | F: ATCCACCCCGCTCTACTCATGTR: TCTCTGTCGTGATGATGTCG |
| 6       | LOC_Os12g23200 | 1,245 | F: GTCTGTCGTGCTGCGTACCA R: GATGAAGGGGTGGCATTGA |
| 7       | LOC_Os06g02490 | 528 | F: AACGAGAAGCAGAGCGCCATC R: GCAGCTTGTTCCTGGTGGAA |

Fig. 4. The similarity of compositions of mRNAs. The abundance of six transcripts (Gene no. 1 to 6) in RNAs extracted with Phenol-SDS, RNeasy and HighGI methods were analyzed by RT-qPCR. Transcript lengths and locus identifiers were listed in Table 2. The quantities were normalized to those of RUBQ2 and shown as relative expression. The same color dots indicate means of technical replicates from the same biological replicate (black, white and grey). Also, means and standard deviation of biological replicates were shown as bars. Asterisks indicate significant differences (paired Student's t-Test, * p < 0.05, ** p < 0.01).
Acknowledgments

This work was supported by JST CREST Grant Number JPMJCR17O1, Japan.

Literature Cited

Barbier, F.F., T.G. Chabikwa, M.U. Ahsan, S.E. Cook, R. Powell, M. Tanurdzic and C.A. Beveridge (2019) A phenol/chloroform-free method to extract nucleic acids from recalcitrant, woody tropical species for gene expression and sequencing. Plant Methods 15: 62.

Chang, S., J. Puryear and J. Cairney (1993) A simple and efficient method for isolating RNA from pine trees. Plant Mol. Biol. Rep. 11: 113–116.

Chirgwin, J.M., A.E. Przybyla, R.J. MacDonald and W.J. Rutter (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18: 5294–5299.

Chomczynski, P. and N. Sacchi (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162: 156–159.

Chomczynski, P., W. Wilfinger, A. Kennedy, M. Rymaszewski and K. Mackey (2010) RNAzol® RT: a new single-step method for isolation of RNA. Nat. Methods 7: 4–5.

Jaakola, L., A.M. Pirittilä, M. Halonen and A. Hohtola (2001) Isolation of high quality RNA from bilberry (Vaccinium myrtillus L.) fruit. Mol. Biotechnol. 19: 201–203.

Kamitani, M., M. Kashima, A. Tetzuka and A.J. Nagano (2019) Lasy-Seq: a high-throughput library preparation method for RNA-Seq and its application in the analysis of plant responses to fluctuating temperatures. Sci. Rep. 9: 7091.

Kawakatsu, T., Y. Wakasa, H. Yasuda and F. Takaiwa (2012) RNA silencing induced by an artificial sequence that prevents proper transcription termination in rice. Plant Physiol. 160: 601–612.

Logemann, J., J. Schell and L. Willmitzer (1987) Improved method for the isolation of RNA from plant tissues. Anal. Biochem. 163: 16–20.

MacDonald, R.J., A.E. Przybyla and W.J. Rutter (1977) Isolation and in vitro translation of the messenger RNA coding for pancreatic amylase. J. Biol. Chem. 252: 5522–5528.

Rohland, N. and D. Reich (2012) Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Res. 22: 939–946.

Sultan, M., V. Amstislavskiy, T. Risch, M. Schuette, S. Dökel, M. Ralser, D. Balzereit, H. Lehrach and M.L. Yaspo (2014) Influence of RNA extraction methods and library selection schemes on RNA-seq data. BMC Genomics 15: 675.

Townley, B.T., M.F. Covington, Y. Ichihashi, K. Zumstein and N.R. Sinha (2015) BrAD-seq: Breath Adapter Directional sequencing: a streamlined, ultra-simple and fast library preparation protocol for strand specific mRNA library construction. Front. Plant Sci. 6: 366.

Ullrich, A., J. Shine, J. Chirgwin, R. Pictet, E. Tischer, W.J. Rutter and H.M. Goodman (1977) Rat insulin genes: construction of plasmids containing the coding sequences. Science 196: 1313–1319.

Yang, H.W., T. Akagi, T. Kawakatsu and R. Tao (2019) Gene networks orchestrated by MeGI: a single-factor mechanism underlying sex determination in persimmon. Plant J. 98: 97–111.

Zeng, Y. and T. Yang (2002) RNA isolation from highly viscous samples rich in polyphenols and polysaccharides. Plant Mol. Biol. Rep. 20: 417.