

Research Note

Isolation of high quality RNA from soil-grown *Ilex paraguariensis* roots suitable for next-generation sequencing and gene expression analysesEdgardo H. Avico¹, Raúl M. Acevedo¹, Pablo I. Calzadilla², Oscar A. Ruiz², Pedro A. Sansberro^{1,*}¹Laboratorio de Biotecnología Aplicada y Genómica Funcional. Instituto de Botánica del Nordeste (CONICET), Facultad de Ciencias Agrarias, Universidad Nacional del Nordeste, Sargento Cabral 2131, CC: 209, W3402BKG, Corrientes, Argentina²Unidad de Biotecnología 1. Instituto de Investigaciones Biotecnológicas-Instituto Tecnológico de Chascomús (UNSAM-CONICET). Av. Intendente Marino Km 8,200 CC: 164. B7130IWA, Chascomús, Buenos Aires, Argentina

*Corresponding author: sansber@agr.unne.edu.ar

Abstract

Extraction of high quality RNA is a prerequisite for downstream application in functional genomics analyses. However, the extraction and purification of pure nucleic acids from root tissues is generally difficult due to the high concentration of carbohydrates and secondary metabolites. Furthermore, the presence of enzymatic inhibitors such as fulvic and humic acids can also negatively affect extraction quality, when extracting from clay soil-grown roots. In this work, total RNA was extracted from soil-grown roots of *Ilex paraguariensis* using four commercially available kits: Spectrum™, RNeasy®, TRI Reagent®, and SV Total RNA Isolation System. Spectral measurement and electrophoresis were used to demonstrate RNA quality and quantity. The Spectrum™ and RNeasy® protocols provided the highest quantity and quality of RNA; however, the former revealed superior extraction performance. Consequently, total RNA was extracted from the roots of non-stressed and drought-stressed plants using the Spectrum™ method and six RNA-seq libraries were prepared from polyA + mRNAs by means of TruSeq mRNA library construction protocol to convert RNA to complementary DNA (cDNA). More than 80 million raw read sequences were obtained from each condition with an average read length of 150 bp. The yield and quality of the total RNA were consistently high and the RNA could be used for further analyses as demonstrated by cDNA library construction, RT-PCR, and transcriptome sequencing. Thus, Spectrum™ method can be used to isolate high quality RNA from roots of normal and drought stressed *I. paraguariensis* plants.

Keywords: Clay soil; dehydration; drought; *Ilex paraguariensis*; RNA isolation; RNA-seq; root tissues, woody plant.**Abbreviations:** PCA_principal component analysis; qRT-PCR_quantitative real-time PCR; RIN_RNA integrity number;

Introduction

Ilex paraguariensis St. Hil. (Aquifoliaceae) is a shade-tolerant tree cultured in northeastern Argentina, southeastern Brazil, and Paraguay. Its leaves and young shoots are used to prepare a tea-like infusion which has stimulant properties and various medicinal benefits (Heck and De Mejia, 2007). Drought and high temperatures during the growing season impose stress on this species, to which it acclimates by triggering a cascade of transcriptomic responses that modify cell metabolism (Acevedo et al., 2016). Stress conditions can affect crop yield, gene expression studies, as well as gene identification and functional characterization based on RNA sequencing, which are important tools for plant breeding.

Current high-throughput genomics analysis requires the isolation of high-quality RNA with high integrity and free of proteins, genomic DNA, and enzyme inhibitors. Integrity is particularly important for quantitative real-time PCR (qRT-PCR) analysis, while purity is essential for reproducibility (Fleige and Pfaffl, 2006). Unfortunately, the presence of carbohydrates and secondary metabolites can interfere with isolation (Ghawana et al., 2011). In effect, tissue of woody plant contains high level of polyphenols, flavonoids and other

compounds which bind and precipitate with nucleic acids. This situation becomes more problematic in case of stressed-tissues, in which a high level of secondary metabolites accumulates as part of the plant defense mechanism to tolerate the adverse environmental conditions (Shinozaki et al., 2015). For this reason, the isolation of suitable RNA from trees is considered more difficult than it is for herbaceous plants (Morante-Carriel et al., 2014). In particular, *I. paraguariensis* tissues present high levels of polyphenols and other secondary metabolites (Filip et al., 2001). Moreover, isolation from root extracts is even more difficult due to their high fiber content, elevated concentration of storage compounds, enzyme inhibitors and contaminants in the soil, in which the roots are grown (Das et al., 2013). Additionally, some soil substances such as humic and fulvic acids can strongly inhibit DNA polymerase (Matheson et al., 2010). To the best of our knowledge, no method has been established for the effective isolation of total RNA from the roots of soil-grown *Ilex paraguariensis* plant. Our aim was to evaluate four commercially available kits, specifically the Spectrum™ Plant Total RNA Kit (Sigma), RNeasy® Plant Mini Kit

(Qiagen), TRI Reagent® (Molecular Research Center), and SV Total RNA Isolation System (Promega), for their ability to extract total RNA from roots. The Spectrum™ kit showed the highest RNA yield and quality among the four, which was tested on roots from non-stressed and drought-stressed plants. Quality of the isolated RNA for downstream application was analyzed by determining the RNA integrity number.

Results and discussion

Comparison of RNA extraction kits

Four RNA extraction kits were evaluated on *Ilex paraguariensis* roots to identify the best method for isolating a large quantity of high quality RNA for transcriptome sequencing. The quality of total RNA analyzed by agarose gel electrophoresis is presented in Fig 1. Intact bands corresponding to 28S and 18S rRNA were clearly visible in RNA samples obtained from TRI Reagent®, Spectrum™ and RNeasy® but not from SV kit (Fig. 2). The sharpness of the ribosomal RNA bands indicated that there was no RNA degradation. Bands of lower molecular weight were also observed corresponding to 5S rRNA. The samples provided different band intensities, indicating differences in total RNA quantity. No DNA contamination was visible in any of the samples.

Concentration and purity of the RNA samples calculated from A₂₆₀, A₂₈₀ and A₂₃₀ are shown in Fig 3. Although, all tested kits yielded detectable RNA from well-watered roots, but the Spectrum™ method provided the highest total RNA, ranged from 53.9 to 75.9 ng·mg⁻¹ DW, whereas RNeasy® and TRI Reagent methods produced lower amounts, ranged from 44.2±16.4 to 22.4 ± 3.8 ng·mg⁻¹ DW, respectively (Fig. 3a). The SV kit was ineffective for root RNA isolation and this result complies with those obtained previously in other recalcitrant species (Zamboni et al., 2008). The absorbance ratio A₂₆₀/A₂₈₀ is generally used to determine protein contamination of a nucleic acid sample and its expected value should be somewhere around 2.1 for high quality RNA with minimum impurities (Gallagher and Desjardins, 2006). In this way, the ratios of isolated RNA from Spectrum™ and RNeasy® kits were in the same ranges and close to the optimum value indicating trace amounts of proteins (Fig. 3b). Meanwhile, the absorbance ratio A₂₆₀/A₂₃₀ indicates the presence of organic contaminants, such as secondary metabolites, soil substances, and reagents (Reck et al., 2015). Samples with A₂₆₀/A₂₃₀ ratios below 1.8 are considered to have a significant amount of these contaminants that will interfere with downstream applications (Sambrook, 2001). Our results have shown that the RNA isolated with commercial kits with included bind-wash-elution methods (Spectrum™, RNeasy®, and SV) had lower impurities (Fig. 3c). In contrast, the direct extraction method using phenol and guanidine isothiocyanate liquid phase separation (TRI Reagent) was less effective to control the impurities present in the sample since the A₂₆₀/A₂₃₀ ratio was lesser than 0.5. Finally, yield and quality were analyzed with PCA to test the efficiency of RNA extraction kits on root tissues harvested from soil-grown plant. Fig 3d shows that the Spectrum™ kit has the best performance like combined yield and purity. Additionally, to evaluate possible downstream applications, we performed RT-PCR analysis of the β -tubulin gene of *I. paraguariensis* (Acevedo et al., 2013). RNA samples obtained from Spectrum™, RNeasy® and TRI Reagent amplified the β -tubulin gene (data not shown).

De novo transcriptome sequencing

According with our previous results, the Spectrum™ protocol was chosen to evaluate the RNA quality for transcriptome sequencing of non-stressed and drought-stressed roots. RIN calculated from the profiles demonstrated the quality assessments of the RNA for subsequent de novo transcriptome sequencing (Fig. 4). The RIN analysis indicated that the integrity of RNA from drought-stressed plants (RIN= 7.2, 7.7, 8.1 out of 10) were lower than those from non-stressed samples (RIN = 8.1, 8.1, 8.6 out of 10). This data reflects that RNA isolated from roots subjected to drought might be more susceptible to RNA degradation than those isolated from non-stressed plants. However, de novo sequencing of RNA samples from roots grown under water deficit and well-watered conditions yielded a total of 87.5 and 84.7 million raw read sequences from 3 libraries, respectively (Fig. 5). The average length of each paired-end read was 150 bp. This high-throughput data generated in this study can later be used to assemble a *de novo* transcriptome using Trinity software (Grabherr et al., 2011).

Materials and methods

Plant material and sample collections

Two-year-old *Ilex paraguariensis* plants were grown in 4 L pots filled with lateritic soil and subjected to controlled environmental conditions (27±1 / 22±2 °C, day / night; 14 h photoperiod; 420 μ mol·m⁻²·s⁻¹ PPFD; 50–55 % RH). They watered regularly during the acclimatization period. Thereafter, water was withheld until the soil water potential (Ψ_{soil}) reached -2 MPa to induce drought stress. Both non-stressed (Ψ_{soil} = -0.04 MPa) and drought-stressed plants were uprooted carefully for the collection of roots. The collected samples were rinsed with tap water, immediately immersed in liquid nitrogen and stored at -70 °C until further processing.

Extraction and purification of total RNA

Total RNA was extracted from the adventitious roots of non-stressed plants by means of the commercial kits Spectrum™ Plant Total RNA Kit (Spectrum™, Sigma), RNeasy® Plant Mini Kit (RNeasy®, Qiagen), TRI Reagent® (TRI Reagent®, Molecular Research Center), and SV Total RNA Isolation System (SV, Promega). Around 100-150 mg root tissue was ground to a fine powder in liquid nitrogen. Extractions were performed by triplicate according to each manufacturer's protocol (Fig. 1). For Spectrum™, RNeasy® and TRI Reagent®, DNA contamination was avoided by including a digestion step with DNase (Invitrogen™ RNase-free DNase I, Thermo Fisher Scientific). Total RNA integrity was verified by resolving 10 μ l of each extraction on a native 1.2 % (w/v) agarose gel stained with ethidium bromide. The bands were visualized under UV light and the images were documented using the gel documentation unit GelDoc-It® (UVP). Meanwhile, total RNA purity was assessed by determining the absorbance of the samples at 230, 260, 280, and 320 nm (background absorption) using a spectrophotometer (SmartSpec™ Plus, Bio-Rad). Extractions with an absorbance ratio A₂₆₀/A₂₈₀ higher than 1.8 were quantified spectrophotometrically at 260 nm (1 OD = 40 ng μ l⁻¹), while samples with an absorbance ratio A₂₆₀/A₂₈₀ lower than 1.8 were estimated by applying the following correction (Glase 1995):

$$\% \text{ Nucleic acid} = (11.16[A_{260}/A_{280}] - 6.32)/(2.16 - [A_{260}/A_{280}])$$

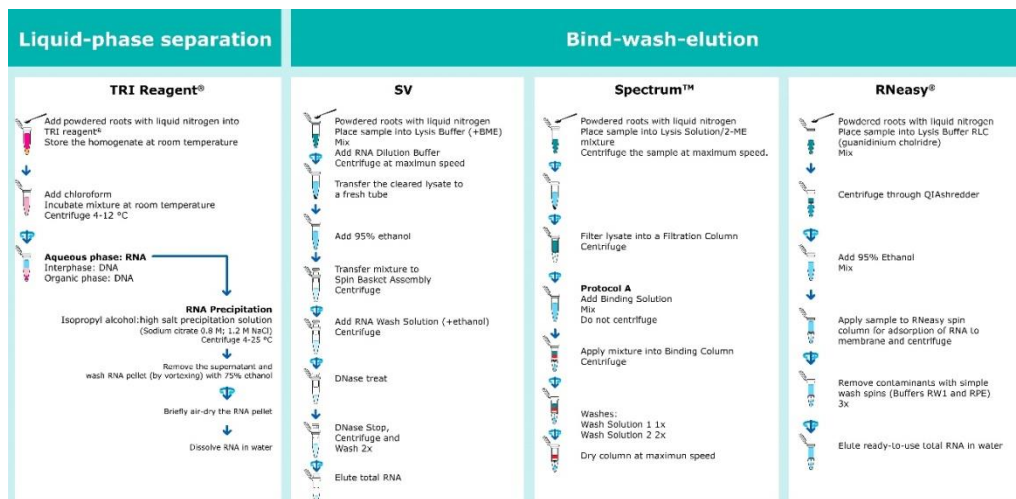


Fig 1. Schematic representation of the protocols used for the isolation of root RNA extracted from soil-grown plants. For Spectrum™, RNeasy® and TRI Reagent®, DNA contamination was avoided by including a digestion step with DNase as described in materials and methods.

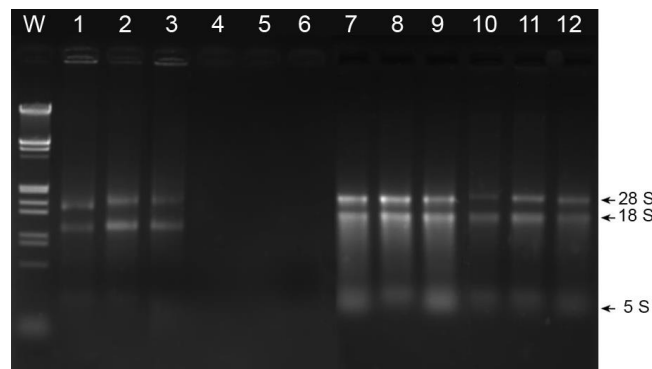


Fig 2. Gel electrophoresis of RNA isolated from roots using four kits. TRI Reagent® (lines 1-3), SV (lines 4-6), Spectrum™ (lines 7-9), and RNeasy® (lines 10-12) kits.

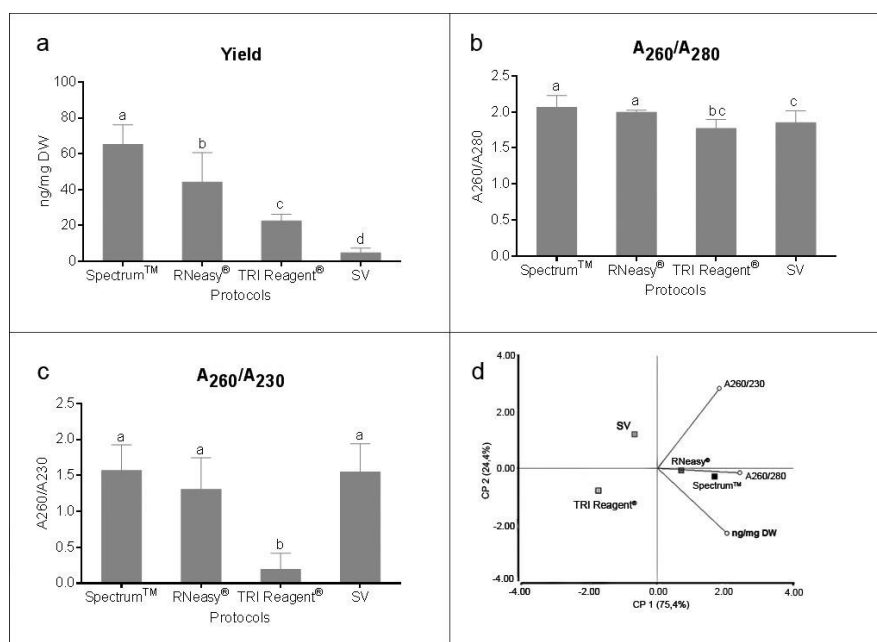


Fig 3. Quantity and quality of root RNA isolated from *I. paraguariensis* using four methods. (a) Total RNA yields. (b-c) Absorbance ratios A_{260}/A_{280} and A_{260}/A_{230} indicate presence of proteins and secondary metabolites impurities, respectively. (d) Principal Components Analysis Biplot relating quality and yield. Means with similar letters do not differ significantly ($P = 0.01$).

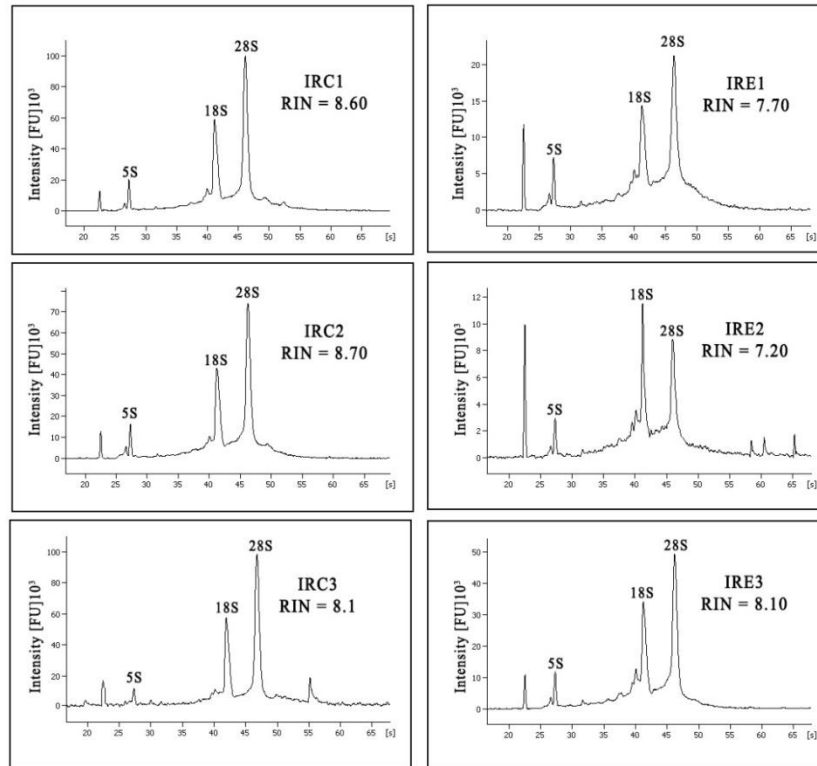


Fig 4. Electropherograms of RNA isolated from root tissues. RNA was extracted from non-stressed (left) and drought-stressed (right) plants with Spectrum™ Plant Total RNA kit. Three biological replicates were used for each treatment. RNA integrity number (RIN) values are indicated within the graph. The 28S and 18S and 5S peaks were marked.

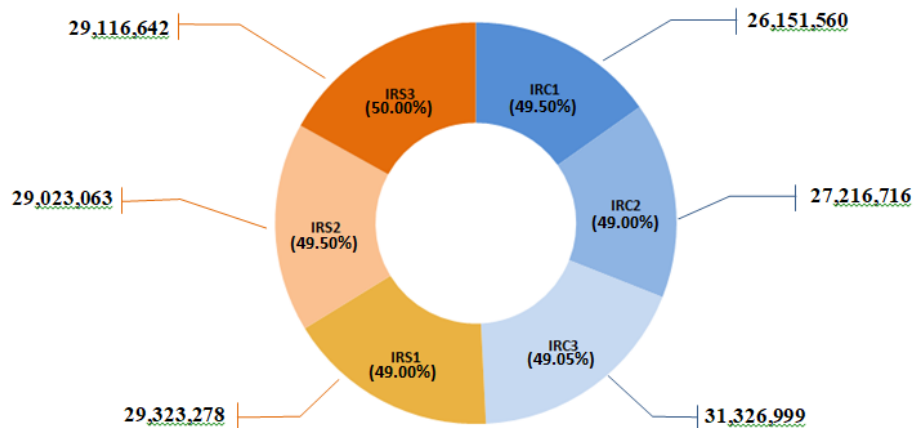


Fig 5. *De novo* transcriptome sequencing of RNA isolated from *I. paraguariensis* roots. Number of paired-ends reads from plants subjected to non-stressed (IRC) and drought-stressed (IRS) conditions. Raw reads data represented the total read count after adaptor eliminations (filtered). Cytosine-guanine contents are shown in parentheses. Total RNA was extracted by using Spectrum™ Plant Total RNA kit.

cDNA synthesis and real-time PCR analysis

First-strand cDNA was synthesized from 1 µg of total RNA from each sample according to the protocol of SuperScript® II Reverse Transcriptase (Thermo Fisher Scientific). Quantitative RT-PCR reactions were prepared with SYBR Green PCR Master Mix (Applied Biosystems). The *I. paraguariensis* β-tubulin gene with the primer sequences 5'-CTGGAGAAGGGATGGATGAA-3' (forward) and

5'-GGCATCCTGATATTGCTGGT-3' (reverse) was amplified as the reference gene (Acevedo et al. 2013). qRT-PCR was carried out on an Applied Biosystems 7500 Real-Time PCR System using SYBR Premix according to the protocol. The thermocycler program consisted of an initial denaturation (95 °C for 10 min), followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. Two biological replicates were amplified in triplicate.

RIN analysis and RNA-sequencing

RIN analysis for RNA-sequencing was achieved from RNA isolations using the Spectrum™ kit from either non-stressed and/or drought stressed plants. Quality control was performed by the system Agilent 2100 Bioanalyzer with a RNA pico/nano Chip platform (Agilent Technologies) and evaluated by interpretation of the RIN value in the range 1–10 (fully degraded RNA – intact RNA).

Six RNA-seq libraries were prepared from polyA+ mRNAs using TruSeq mRNA library construction protocol (Illumina). High-throughput pair-end (2x150bp) sequencing was performed on a HiSeq 1500 platform (Illumina). The quality of raw read data was accessed by FastQC program (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>). Trimmomatic program (Bolger et al. 2014) was used to remove adapters from the raw reads.

Experimental design and statistical analysis

The experiment was repeated three times by performing three biological units. The data sets were subjected to the Shapiro-Wilk test for normality ($\alpha = 0.05$) and were statistically analyzed by ANOVA (GraphPad version 7.0). Tukey's multiple comparison test was used to compare differences among the treatments. A value of $P = 0.01$ was used to assess statistical significance. Principal component analysis (PCA) was performed with Infostat®.

Conclusion

The Spectrum™ Plant Total RNA kit (Sigma) was demonstrated to be the most effective method for extracting a large amount of high quality RNA required for transcriptome sequencing of *I. paraguariensis* roots, even from water-stressed plants.

Acknowledgements

This work was supported by grants from Agencia Nacional de Promoción Científica y Tecnológica (PICT 2014-1246), Secretaría General de Ciencia y Técnica- Universidad Nacional del Nordeste (PI A001/14), and Establecimiento Las Marías S.A. (Gdor. Virasoro, Corrientes, Argentina). OAR and PAS are members of the Research Council of Argentina (CONICET). EHA and PIC received CONICET scholarships. RMA is Professor of the Universidad Nacional del Nordeste (UNNE).

Conflict of Interest

The authors declare that they have no conflict of interest.

Author contributions

Conceived and designed the experiments: EHA, OAR and PAS. Performed the experiments: EHA, MHA and PIC. Analyzed the data: EHA, MHA and PAS. Written the manuscript: PAS.

References

- Acevedo RM, Maiale SJ, Pessino SC, Bottini R, Ruiz O, Sansberro P (2013) A succinate dehydrogenase flavoprotein subunit-like transcript is upregulated in *Ilex paraguariensis* leaves in response to water deficit and abscisic acid. *Plant Physiol Bioch.* 65:48–54.
- Acevedo RM, Ruiz O, Sansberro P (2016) Transcriptome expression profiling in response to drought stress in *Ilex paraguariensis* leaves. *Plant Omics.* 9:334–343.
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 30:2114–20.
- Das A, Saha D, Mondal TK (2013) An optimized method for extraction of RNA from tea roots for functional genomics analysis. *Indian J Biotechnol.* 12:129–132.
- Filip R, Lopez P, Giberti G, Coussio J, Ferraro G (2001) Phenolic compounds in seven South American *Ilex* species. *Fitoterapia.* 72:774–778.
- Fleige S, Pfaffl MW (2006) RNA integrity and the effect on the real-time qRT-PCR performance. *Mol Aspects Med.* 27:126–139.
- Gallagher SR, Desjardins PR (2008) Quantitation of DNA and RNA with absorption and fluorescence spectroscopy. *Curr Protoc Protein Sc.* 76:3D:A.3D.1–A.3D.21.
- Ghawana S, Paul A, Kumar H, Kumar A, Singh H, Bhardwaj PK, Rani A, Singh RS, Raizada J, Singh K and Kumar S (2011) An RNA isolation system for plant tissues rich in secondary metabolites. *BMC Res Notes.* 4:85–89.
- Glasel JA (1995) Validity of nucleic acid purities monitored by 260nm/280nm absorbance ratios. *Biotechniques.* 18:62–63.
- Heck CI, De Mejia EG (2007) Yerba mate tea (*Ilex paraguariensis*): a comprehensive review on chemistry, health implications, and technological considerations. *J Food Sci.* 72:138–151.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A (2011) Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat Biotechnol.* 29:644–52.
- Matheson CD, Gurney C, Esau N, Lehto R (2010) Assessing PCR inhibition from humic substances. *Open Enzym Inhib J.* 3:38–45.
- Morante-Carriel J, Sellés-Marchart S, Martínez-Márquez A, Martínez-Esteso MJ, Luque I, Bru-Martínez R (2014) RNA isolation from loquat and other recalcitrant woody plants with high quality and yield. *Anal Biochem.* 452:46–53.
- Reck M, Tomasch J, Deng Z, Jarek M, Husemann P, Wagner Döbler I (2015) Stool metatranscriptomics: a technical guideline for mRNA stabilisation and isolation. *BMC Genomics.* 16:1.
- Sambrook JR (2001) Molecular cloning. A laboratory manual, Third ed. CSHL Press, New York.
- Shinozaki K, Uemura M, Bailey-Serres J, Bray EA, Weretilnyk E (2015) Plant responses to abiotic stress. In: Buchanan BB, Gruissem W, Russell LJ (ed) *Biochemistry and molecular biology of plants*, 2nd edn. Wiley, Chichester, UK. Pp:1051–1100.
- Zamboni A, Pierantoni L, De Franceschi P (2008) Total RNA extraction from strawberry tree (*Arbutus unedo*) and several other woody-plants. *IFOREST.* 1:122–125.