Optimization of insect genomic DNA and total RNA extraction protocols for high fidelity gene sequencing

Eliza F. de Melo Bellard do Nascimento
Embrapa Genetic Resources and Biotechnology https://orcid.org/0000-0002-8140-8168
Vívian dos Santos Lucena-Leandro
Embrapa Genetic Resources and Biotechnology https://orcid.org/0000-0003-1497-9144
Caroline R. Torres
Brasilia University https://orcid.org/0000-0001-9677-6484
Leonardo A. Vidal
Brasilia University https://orcid.org/0000-0003-3259-3528
Camila Ivo C. V. F. Junqueira
Brasilia University https://orcid.org/0000-0001-9126-2724
Juliana Dantas de Almeida
Embrapa Genetic Resources and Biotechnology https://orcid.org/0000-0002-2422-5808
Érika Valéria Albuquerque Saliba Albuquerque (erika.albuquerque@embrapa.br)
Embrapa Genetic Resources and Biotechnology https://orcid.org/0000-0003-3967-2384

Method Article

Keywords: DNA, Illumina, insect, PacBio, RNA, sequencing.

Posted Date: October 11th, 2022

DOI: https://doi.org/10.21203/rs.3.pex-1928/v1

Abstract

There is a strong demand for sustainable, durable, safe, and specific solutions against agricultural pests due to breeding limitations and chemical control constraints. Biotechnological approaches to insect pests require phylogenetic information at the molecular level for RNAi application and other tailored strategies, such as insecticide resistance molecular markers. Sequencing technologies have advanced in resolution, speed, and cost-effectiveness to offer striking impact on genomics and transcriptomics. To provide deep and reliable data to high-throughput studies, state-of-the art technologies require high purification and integrity of the submitted samples. However, the limited number of existing protocols to obtain high mass nucleic acids from insects render difficult to obtain elevated purity and integrity standards. Species characteristics, as scales covered bodies of Lepidoptera and mining eating habits, hinder the quality of nucleic acid samples. Additionally, small size individuals are limited to yield sufficient amounts to robust base sequence coverage. The coffee leaf miner (CLM) *Leucoptera coffeella* is microlepdoptera insect responsible for severe leaf damage and huge yield losses to the coffee crop. Here we present simple and reproducible procedures we used to successfully obtain the whole-genome long-read PACBIO and Illumina (DNA) and transcriptome (RNA) of *L. coffeella* for *de novo* assembly. Therefore, we developed customized protocols applicable to insects to the extraction of genomic High Molecular Weight DNA and total RNA suitable for downstream high standard sequencing applications and other sensitive molecular analyses.

Introduction

Among the insects of economic importance are the miners, which feed on the mesophyll of agricultural species leaves. The most remote occurrence records of miner insects date back to the Triassic, the first mesozoic period, approximately 250 million years ago1. In this group, flies such as *Liriomyza bryoniae* (Tomato Leaf Miner) 2,3, beetles such as *Taphrocerus cocois* (Coconut Leaf Miner)4 and moths such as *Leucoptera coffeella* (Coffee Leaf Miner)5 are highlighted. They are phytophagous insects whose damage to the target host is related to the way they feed in the larval stage. In this case, the larva feeds on the region amidst the two epidermis of the leaves forming tunnels here called mines, causing necrosis and defoliation, which compromise the photosynthetic capacity of the plant, affecting its viability 6–8.

The phytosanitary department of the European Food Safe Authority (ESFA), published in 2020 a panel concerning species threats and among them there are mining insects 3. Although being economically relevant in several production chains, miner insects aren't widely studied due to their cryptic habits9, which makes it difficult to obtain adequate samples for phylogenetic, epidemiological and phytosanitary studies. An initiative to sequencing 100 pest genomes has generated genomic data from arthropod pests of agricultural species. The selection of insects followed predetermined criteria. One of them was the

existence of high-quality nucleic acid extraction methods, where the miner insects were not contemplated 10.

Our group has been dedicated to studying the Coffee Leaf Miner (CLM) (*Leucoptera coffeella*) (Guérin-Mèneville & Perrottet) (Lepidoptera: Lyonetiidae) from several approaches such as life cycle, morphological features and, more recently, the genome and transcriptome. CLM is an exclusive pest of coffee trees where it could cause losses up to 87% in yield 11,12. Coffee is grown by millions of producers in more than 60 countries, which reflects in a wide production chain, considered one of the most profitable in the world, which is threatened by pests such as CLM 13,14. The control of CLM with chemical, pesticides is necessary, however it entails disadvantages such as the selection of resistant populations, high costs of production and harmful effects on the environment and human health 8,15,16. For this reason, it is important to develop strategies for the implementation of integrated pest management systems (IPM), which balance the increase and quality of grain production and the protection of the environment 8,17.

In this way, the molecular characterization of CLM with the sequencing of its reference genome and transcriptome, are required for identification of genetic based tools. Moreover, those data will allow further studies such as the exploration of target genes for gene silencing via interfering RNA (RNAi) and molecular markers. These approaches will enable the development of biotechnological solutions aimed at controlling this important pest in coffee parks. The genomic sequencing technique chosen for this approach was PacBio technology, which allows the reading of long fragments, reaching up to 60–80 kb 18. However, genomic DNA (gDNA) samples of high molecular weight, which allows the reading of longer lengths were not obtained by methods already described due to the high chitin content in the structure of CLM and also to the feeding habits that impairs the extraction. These restrictions also limit obtaining high quality RNA samples. Here we present a new method of extraction of high quality and repeatable total gDNA and RNA, from different stages of development of CLM for Long-Read sequencing (PacBio) and MiSeq Illumina, respectively.

Reagents

Kit E.Z.N.A insect DNA (Omega BioTek, Cat. No.: D0926-02)

Kit ReliaPrep RNA Tissue Miniprep System (Promega, Cat.No.: PRZ6111).

Proteinase K (Invitrogen, Cat. No.: AM2544)

Chloroform (Sigma-Aldrich, Cat. No.: C2432)

Isoamyl alcohol (Sigma-Aldrich, Cat. No.: W205702)

RNAse A (Invitrogen, Cat. No.: 12091021)

Ethyl alcohol, Pure (Sigma-Aldrich, Cat. No.: E7023)

TRI Reagent® (Invitrogen, Cat.No.: 15596026).

Isopropyl alcohol (Sigma-Aldrich, Cat. No.: I9516)

Equipment

Incubator Shaker Ecotron (Infors HT)

Refrigerated Centrifuge (Eppendorf - 5415R)

MultiTemp®III (PharmaciaBiotech)

NanoDrop 2000/2000c Spectrophotometer (Thermo Fisher Scientific)

Qubit Fluorometric Quantification (Thermo Fisher Scientific)

Femto Pulse (Agilent)

Bioanalyzer 2100 system (Agilent)

Procedure

Genomic DNA extraction

Extractions were performed from CLM pupae collected in *Coffea arabica* leaves during the dry and hot season, from April to September 2021, in the brazilian savanna region of the central plateau (experimental field of Embrapa Cerrados, Planaltina - DF, latitude: 15° 35' 30' 'S, longitude 47° 42' 30" W).

The E.Z.N.A insect DNA kit (Omega BioTek, Cat. No.: D0926-02) was used to obtain the insect gDNA. However, considerable modifications of the protocol provided by the kit was required in order to obtain a higher amount and quality of gDNA extracted.

1- Cell lysis: The standard kit protocol provided low yield for our samples. Therefore, in order to increase the yield of the samples we suggest: macerate 20 pupae (around 10 mg) in 700 uL of lysis buffer (CTL) for 2 minutes in a *Dounce* tissue grinder (or homogenizer) instead at liquid nitrogen as recommended by the kit manufacturer. Transfer the lysate to a sterile 1.5 mL microcentrifuge tube with the aid of a 1 mL automatic pipette with a wide terminus tip.

2- Protein Digestion: add 100 mg of proteinase K (Invitrogen, Cat. No.: *AM2544*) to the tube containing the macerated tissue in CTL buffer, slowly invert 10 times and incubate for 16 hours at 37 °C under mild agitation (100 rpm) (Incubator Shaker Ecotron - Infors HT).

3- DNA extraction: add 700 µL of chloroform:isoamyl alcohol (24:1) to the microtube from the previous step and carefully invert 20 times. In order to avoid genome DNA shearing, the use of homogenizers like vortex are not recommended. Centrifuge at a 10,000 g speed in an *Eppendorf*® 5415R Microcentrifuge apparatus for 2 minutes, room temperature (RT). After centrifugation, transfer approximately 500uL of the aqueous phase (upper one) to a 2 mL fresh tube using a 1 mL automatic pipette with a wide terminus tip. Cellular debris deposited at the interface between the polar and apolar phases should be discarded.

4- RNAse digestion: add 0,06 ng of RNAse A (Invitrogen, Cat. No.: 12091021) to the microtube from the previous step containing the aqueous phase. Carefully invert 10 times and keep at 37 °C for 30 minutes. Add 500 µL of BL buffer and carefully invert the tube 20 times and incubate for 10 minutes 70 °C.

5- DNA Precipitation: add 1 µL of ethanol 100% to the microtube from the previous step and mix by inverting the tube slowly 20 times. In order to avoid genome DNA shearing the use of homogenizers like vortex are not recommended.

6- DNA Purification: for DNA binding, load 750 µL of the microtube content of the previous step on the top of a Mini Column HiBand DNA already placed in a collect tube of 2 mL (column and tube provided by the kit) with the aid of a 1 µL automatic pipette with a wide terminus tip. Centrifuge the set of column plus collect tube in a 15,000 g speed for 30 seconds and discard the flow-through and repeat this step until sample is over. Carry on the last centrifugation by 1 minute. Clean the DNA by adding 500 µL of HBC buffer. Centrifuge at a 15,000 g speed for 1 minute. Discard the flow-through and load 700 µL of wash buffer (provided by the kit). Repeat the centrifugation and cleaning for a best result. Place the column in a fresh microtube.

7- DNA elution: For DNA elution, centrifuge the column where the DNA was bound in a 15,000 g speed during 2 minutes to eliminate any trace of binding buffer and place in a fresh 1,5 mL microtube. Carefully load 15 uL of the elution buffer (provided by the kit) on the middle of the column membrane. Incubate at RT for 5 min, and after centrifuge at a 15,000 g speed for 30 s to collect the eluted DNA from the column. Repeat the elution step to maximum yield.

8- Quality control and storage: evaluate the eluted DNA for its concentration and quality in devices such as NanoDrop, Qubit, agarose gel electrophoresis and Femto Pulser. The samples must be stored at 4°C. If necessary, the samples can be frozen at -20°C, however the quality might be compromised.

The main steps of gDNA extraction are illustrated in Figure 1.

Total RNA extraction

The extractions were performed of 7 stages of development from CLM: 4 larval instars (L1, L2, L3 and L4) (Motta et al., 2021), pupae, males and females. TRIzol® method was used, followed by RNA cleanup with ReliaPrep RNA Tissue Miniprep System kit (Promega, Cat.No.: PRZ6111).

1- Cell lysis: with the aid of a mortar and pestle, macerate separately the different stages of development (approximately 20 µL of powder per tube) with liquid nitrogen. Dissolve the material in a sterile 1,5 mL microfuge tube containing 1 mL of TRIzol® Reagent (Invitrogen, Cat.No.: 15596026). Invert the tube carefully 10 times and incubate it for 5 minutes at RT.

2- Organic extraction: centrifuge the tube at 10,000 g speed (Refrigerated Centrifuge Eppendorf - 5415R) for 10 minutes at 4 °C for precipitation of impurities and other undissolved materials. Transfer the supernatant to a new sterile 1,5 mL microfuge tube containing 200 mL of chloroform. Agitate the tube vigorously by manual inversion for 15 s. Incubate for 3 minutes at RT and centrifuge at 10,000 g speed for 15 minutes at 4°C.

3- RNA precipitation: carefully transfer the aqueous phase (upper one) from the previous step to a new microtube containing 500 µL of isopropanol. Slowly invert the tube 10 times and incubate for 10 minutes at RT. After this, centrifuge RNA at 10,000 g speed for 10 minutes at 4 ºC to sediment and discard the supernatant.

4- RNA cleaning: after obtaining the precipitated, discard the supernatant and add 1 mL of ethanol 75%. Carefully invert 5 times the tube containing the pellet and centrifuge at 10,000 g for 5 minutes at 4 ºC. Discard the Ethanol and dry the tube containing the pellet at RT.

5- RNA purification with ReliaPrep RNA Tissue Miniprep System: dissolve the pellet in 250 µL of LBA/TG solution (supplied by the kit) and add 85 µL of isopropanol. Insert the sample into a ReliaPrep™ Mini Column (supplied by the kit) and centrifuge at 10,000 g for 1 minute at RT. Discard the flowthrough and add 500 µL of RNA Wash Solution on the top of the column. Centrifuge at 10,000 g for 30 s at RT. Meanwhile, prepare a DNAse mix containing 24 µL de Yellow Core Buffer + 3 µL $MnCl_2$ + 3 µL DNAse (reagents supplied by the kit). Add the DNAse mix on the top of the column and centrifuge at 10,000 g for 15 minutes at RT. Add 200 µL of Column Wash solution and centrifuge at 10,000 g for 15 minutes at RT. Add 500 µL of RNA Wash Solution to the column and centrifuge at 10,000 g for 30 s at RT. Discard the flowthrough, add 300 µL of RNA Wash Solution to the column and centrifuge at 10,000 g for 2 minutes at RT. Transfer the column to a fresh microtube, add 15 µL of RNAse free water on the top of the column and centrifuge at 10,000 g for 1 minute at RT.

6- Quality control and storage: after extraction, evaluate the RNA concentration and quality in NanoDrop, Qubit and Bioanalyzer. RNA should be stored immediately in a -80°C freezer to prevent degradation.

The main steps of RNA extraction are illustrated in Figure 2.

Troubleshooting

DNA

Several features of insects such as high amount of chitin in its structure and feeding habits which bring contaminants to the samples disturb the obtaining of high-quality genetic material 19. Some described methods of gDNA extraction used in other insects19 have not been successfully applied to our goal, even those recommended by PacBio such as *Anopheles coluzzii*, *Tribolium castaneum* and *Hyles vespertilio* 18,20,21. After carrying out different tests of extractions of different stages of development (pupae, larval stages and adult), we concluded that pupae provided better quality DNA because they had less chitin in their structure and almost no content in their guts.

Key changes were performed in the E.Z.N.A insect DNA kit protocol: 1) the material was not macerated in liquid nitrogen. To promote cell lysis, pupae were macerated in *Dounce* tubes with CTL buffer (provided by the kit); 2) to avoid gDNA degradation, automatic pipette with a wide terminus tip were used for any manipulation of the sample; 3) The samples were not homogenized by vortexing. Instead, slow movements were performed by inversion; 4) During cell lysis and protein digestion, samples were incubated longer at a lower temperature and under gentle agitation to increase yield and reduce DNA degradation; 5) instead eluting DNA with pre-heated buffer, a higher integrity was revealed using RT e 6) the volume of the elution buffer was reduced to obtain more concentrated samples. These recommendations provided significant improvements for obtaining *L. coffeella* gDNA and may be useful for other insect species.

RNA

In general, the extraction of total RNA for sequencing presents methodological challenges to keep stability of the molecules. RNA is susceptible to degradation since the initial steps with cell lysis until final quantification 22. To avoid degradation and loss of transcriptome information, RNA extraction protocols suggest strict procedures during manipulation, such as the use of exclusive equipment for RNA and isolated and decontaminated areas of RNase 23.

The protocol established here provided for obtaining total RNA from the different developmental stages of *L. coffeella*. In addition to the common steps of conventional protocols, we added the step of purification of the samples extracted using a commercial extraction kit, as a differential of the methodology. In comparison to other RNA extraction methods, this protocol showed a higher degree of purity and concentration of the samples.

Although there are relevant ribosomal RNA (rRNA) profiles for various model organisms 24, there are still few recent studies that assess metrics and quality standards for non-model organisms 25. Most analyzes

of insect RNA profiles document the “hidden denaturation” of the 28S subunit as an apparently degraded profile 26–29. However, with increased sequencing of different organisms for RNA investigations in recent years, more invertebrate rRNA profiles have been generated and suggest that denaturation of the 28S subunit is indeed very common 25,28,29.

Time Taken

DNA

The extraction of gDNA is performed in two days. The first consists of overnight cell lysis and digestion for approximately 16 hours. The second day, which includes the remaining stages, takes approximately 1 hour and 30 min.

RNA

Trizol® RNA extraction takes approximately 1 hour and 45 min. Purification of RNA with the commercial kit takes approximately 45 min.

Anticipated Results

DNA

We present a new, fast and reproducible methodology for obtaining high quality CLM gDNA for Long-Read (PacBio) and Illumina sequencing. The extracted gDNA samples showed concentrations, absorbance ratios (A260/A280 and A260/A230) and integrity recommended for Long-Read sequencing (PacBio). Evaluations of these parameters were conducted in NanoDrop and Qubit (Table 1), agarose gel (Figure 3) and Femto Pulse (Figure 4).

Total RNA

Samples extracted using this protocol were evaluated on NanoDrop and Qubit spectrophotometers and met the specific concentration and ratio criteria (A260/A280 and A260/A230) required for high-throughput transcriptome sequencing (Table 2). RNA integrity was confirmed by agarose gel electrophoresis (Figure 5) and Bioanalyzer (Figure 6), currently the two methods used to evaluate RNA integrity and degradation after extraction. The Bioanalyzer calculates the level of RNA integrity (RNA integrity number - RIN) from the distribution peaks of RNA fragments, which varies from 1 to 10. Values

above 7 are considered high quality and suitable for sequencing [30]. Another way of assessing RNA integrity considers the size of the rRNA subunits: between 4400 and 5000 base pairs (bp) for the 28S subunit and 2000 bp for the 18S subunit [31].

The migration profile of the samples on non-denaturing agarose gel showed an rRNA band of approximately 850 bp (Figure 5) while the electropherogram generated by the Bioanalyzer device shows the presence of a rRNA band with approximately 2000 bp (Figure 6A). This discrepancy in the height of the RNA band can be explained by intramolecular interactions where RNA molecules can bend and change the secondary structure of the molecules, affecting migration in the non-denaturing agarose gel [32]. More accurate size estimates can be obtained using denaturing gel electrophoresis in which hydrogen bonds are broken allowing the RNA to migrate as a single-stranded molecule [32].

At first these results were interpreted as a degradation profile, as traditionally the high quality RNA presentes the subunits 28S and 18S [32]. However, a particularity is described for invertebrates, where the subunit 28S is highly prone to denaturation, resulting in two smaller fragments that overlaps with subunit 18S in a single band [28,29,33], as observed here. Moreover, the distribution peaks of RNAr fragments showed a high RIN, as an indicator of the high quality of the samples (Figure 6B).

The protocol optimization of RNA extraction from CLM contributes to the attainment of biological material highly purified that allows the use of sequencing methods that generate more robust data. The methodology described here allows the surrender of significant results for the generation of knowledge in the areas of genome and transcriptome and can be useful in other molecular sequencing strategies, such as microRNAs, epigenomics, among others. In addition, molecular techniques of PCR, qPCR and others can benefit from these protocols.

More broadly, the application of the data generated from these protocols will enable the targeting of specific and sustainable insect control strategies, such as gene silencing via RNAi and the development of molecular markers of insecticide resistance. Is expected it will contribute to studies of insecticidal bio inputs of species harmful to plants for forest, ornamental, medicinal and agricultural use.

References

1. Imada, Y., Oyama, N., Shinoda, K., Takahashi, H. & Yukawa, H. Oldest leaf mine trace fossil from East Asia provides insight into ancient nutritional flow in a plant–herbivore interaction. *Sci Rep* **12**, 5254 (2022).

2. da COSTA-LIMA, T. C., SILVA, A. de C. & Parra, J. R. P. Moscas-minadoras do gênero Liriomyza (Diptera: Agromyzidae): aspectos taxonômicos e biologia. *Embrapa Semiárido-Documentos (INFOTECA-E)* (2015).

3. Health (PLH), E. P. on P. *et al.* Pest categorisation of Liriomyza bryoniae. *EFSA Journal* **18**, e06038 (2020).

4. Ferreira, J. M. S. & Morin, J. P. Aspéctos da biologia e comportamento do minador da folha do coqueiro, *Taphrocerus cocois* Bondar, 1922 (Coleoptera : Buprestidae) e ocorrência de parasitismo. *Anais da Sociedade Entomológica do Brasil* 3 (1985).

5. Motta, I. O. *et al.* The coffee leaf miner, Leucoptera coffeella (Lepidoptera: Lyonetiidae): identification of the larval instars and description of male and female genitalia. *Revista Brasileira de Entomologia* **65**, (2021).

6. Guérin-Méneville, F.-É. & Perrottet, S. *Mémoire sur un insecte et un champignon qui ravagent les cafiers aux Antilles*. (Bouchard-Huzard, 1842).

7. Costa, J. N. M., Teixeira, C. A. D., Júnior, J. R. V., Rocha, R. B. & Fernandes, C. de F. Informações para facilitar a identificação das diferentes fases do bicho-mineiro (Leucoptera coffeella) em campo. *Embrapa Rondônia-Comunicado Técnico INFOTECA-E, Rondônia* (2012).

8. Dantas, J. *et al.* A Comprehensive Review of the Coffee Leaf Miner Leucoptera coffeella (Lepidoptera: Lyonetiidae)—A Major Pest for the Coffee Crop in Brazil and Others Neotropical Countries. *Insects* **12**, 1130 (2021).

9. Derocles, S. A. P., Evans, D. M., Nichols, P. C., Evans, S. A. & Lunt, D. H. Determining Plant – Leaf Miner – Parasitoid Interactions: A DNA Barcoding Approach. *PLOS ONE* **10**, e0117872 (2015).

10. Childers, A. K. *et al.* The USDA-ARS Ag100Pest Initiative: High-Quality Genome Assemblies for Agricultural Pest Arthropod Research. *Insects* **12**, 626 (2021).

11. Souza, J. C. de, Reis, P. R., Rigitano, R. L. de O. & Ciociola Júnior, A. I. Eficiência de thiamethoxam no controle do bicho-mineiro do cafeeiro. II - Influência na época de aplicação via irrigação por gotejamento. (2006).

12. Neves, M. F. Análise Dos Benefícios Econômicos e Sociais Da Utilização Do Carbofurano No Controle de Nematoides, Bicho Mineiro (Leucoptera coffeella) e Cigarra Do Cafeeiro (Quesada Gigas e Fidicina Pronoe) Na Cultura Do Café. *Univ. Sao Paulo* **29**, 1–29 (2016).

13. Concafe. Produção dos Cafés do Brasil atinge 61,62 milhões de sacas de 60kg em 2020, volume 25% maior que 2019. http://www.consorciopesquisacafe.com.br/index.php/imprensa/noticias/1023-2020-09-29-15-22-05 (2020).

14. USDA, U. S. D. of A. *Coffee: World Markets and Trade*. (2021).

15. Leite, S. A. *et al.* Profile of Coffee Crops and Management of the Neotropical Coffee Leaf Miner, Leucoptera coffeella. *Sustainability* **12**, 8011 (2020).

16. Leite, S. A. *et al.* Time-concentration interplay in insecticide resistance among populations of the Neotropical coffee leaf miner, Leucoptera coffeella. *Agricultural and Forest Entomology* **23**, 232–241

(2021).

17. Sawinska, Z., Świtek, S., Głowicka-Wołoszyn, R. & Kowalczewski, P. Ł. Agricultural Practice in Poland Before and After Mandatory IPM Implementation by the European Union. *Sustainability* **12**, 1107 (2020).

18. Oppert, B., Stoss, S., Monk, A. & Smith, T. Optimized Extraction of Insect Genomic DNA for Long-Read Sequencing. *Methods and Protocols* **2**, 89 (2019).

19. Wouters, R., Mugford, S., Biello, R., Heavens, D. & Hogenhout, S. Extraction of high molecular weight DNA from aphids and other sap-feeding insects for long-read sequencing. *protocols.io* https://www.protocols.io/view/extraction-of-high-molecular-weight-dna-from-aphid-bhftj3nn (2020).

20. Kingan, S. B. *et al.* A High-Quality De novo Genome Assembly from a Single Mosquito Using PacBio Sequencing. *Genes* **10**, 62 (2019).

21. Pippel, M. *et al.* A highly contiguous genome assembly of the bat hawkmoth Hyles vespertilio (Lepidoptera: Sphingidae). *GigaScience* **9**, (2020).

22. Houseley, J. & Tollervey, D. The Many Pathways of RNA Degradation. *Cell* **136**, 763–776 (2009).

23. Florell, S. R. *et al.* Preservation of RNA for Functional Genomic Studies: A Multidisciplinary Tumor Bank Protocol. *Mod Pathol* **14**, 116–128 (2001).

24. Schroeder, A. *et al.* The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC molecular biology* **7**, 1–14 (2006).

25. McCarthy, S. D., Dugon, M. M. & Power, A. M. 'Degraded'RNA profiles in Arthropoda and beyond. *PeerJ* **3**, e1436 (2015).

26. Ishikawa, H. t & Newburgh, R. W. Studies of the thermal conversion of 28 S RNA of Galleria mellonella (L.) to an 18 S product. *Journal of Molecular Biology* **64**, 135–144 (1972).

27. Fujiwara, H. & Ishikawa, H. Molecular mechanism of introduction of the hidden break into the 28S rRNA of insects: implication based on structural studies. *Nucleic acids research* **14**, 6393–6401 (1986).

28. Winnebeck, E. C., Millar, C. D. & Warman, G. R. Why does insect RNA look degraded? *Journal of Insect Science* **10**, 159 (2010).

29. DeLeo, D. M., Pérez-Moreno, J. L., Vázquez-Miranda, H. & Bracken-Grissom, H. D. RNA profile diversity across arthropoda: guidelines, methodological artifacts, and expected outcomes. *Biology Methods and Protocols* **3**, bpy012 (2018).

30. Mueller, O., Lightfoot, S. & Schroeder, A. RNA integrity number (RIN)–standardization of RNA quality control. *Agilent application note, publication* **1**, 1–8 (2004).

31. Melen, G. J., Pesce, C. G., Rossi, M. S. & Kornblihtt, A. R. Novel processing in a mammalian nuclear 28S pre-rRNA: tissue-specific elimination of an 'intron'bearing a hidden break site. *The EMBO Journal* **18**, 3107–3118 (1999).

32. Sambrook, J. & Russell, D. W. Molecular cloning: A laboratory manual. Mol. *Cloning A Lab. Man* (2001).

33. Ogino, K., Eda-Fujiwara, H., Fujiwara, H. & Ishikawa, H. What causes the aphid 28S rRNA to lack the hidden break? *Journal of molecular evolution* **30**, 509–513 (1990).

Acknowledgements

The authors thank the researcher Adriano Veiga from Embrapa Cerrados, for providing access to the coffee plantations and for helping in insect collection and to Genomix Data, by the scientific advisory during the development of work. Work supported by the Consortium Brazilian Research and Development of Coffee - CBP & D Exchange / Coffee No. 10.18.20.004.00.00.

Figures

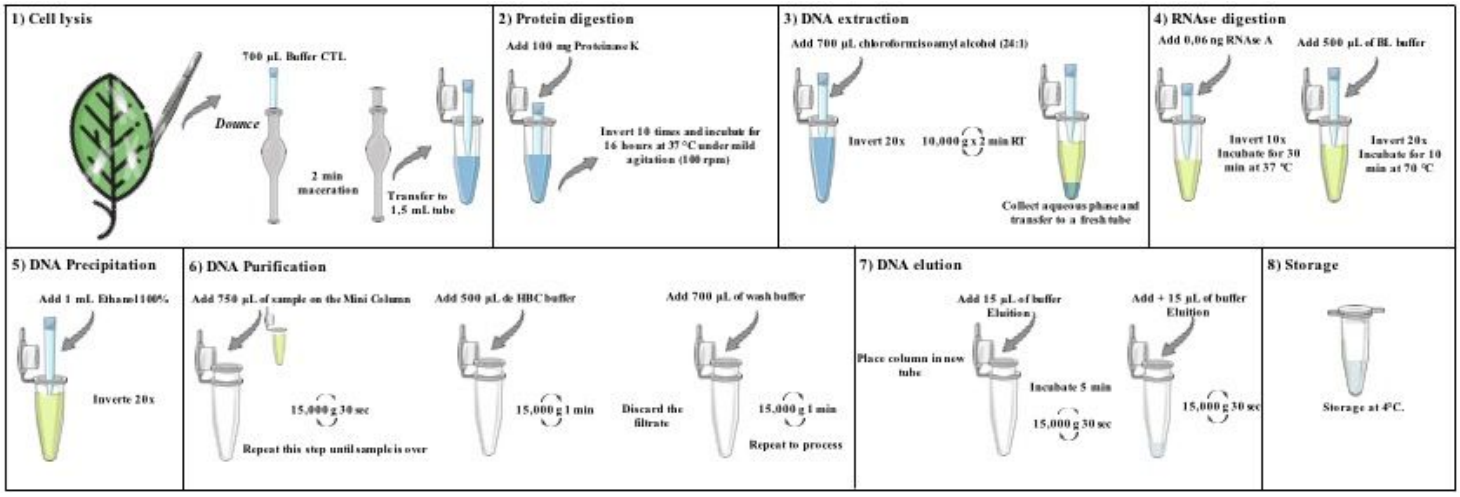

Figure 1

Synthesis of optimized gDNA extraction steps: 1) Cell lysis; 2) Protein Digestion; 3) DNA extraction; 4) RNAse digestion; 5) DNA Precipitation; 6) DNA Purification; 7) DNA elution ans 8) Storage.

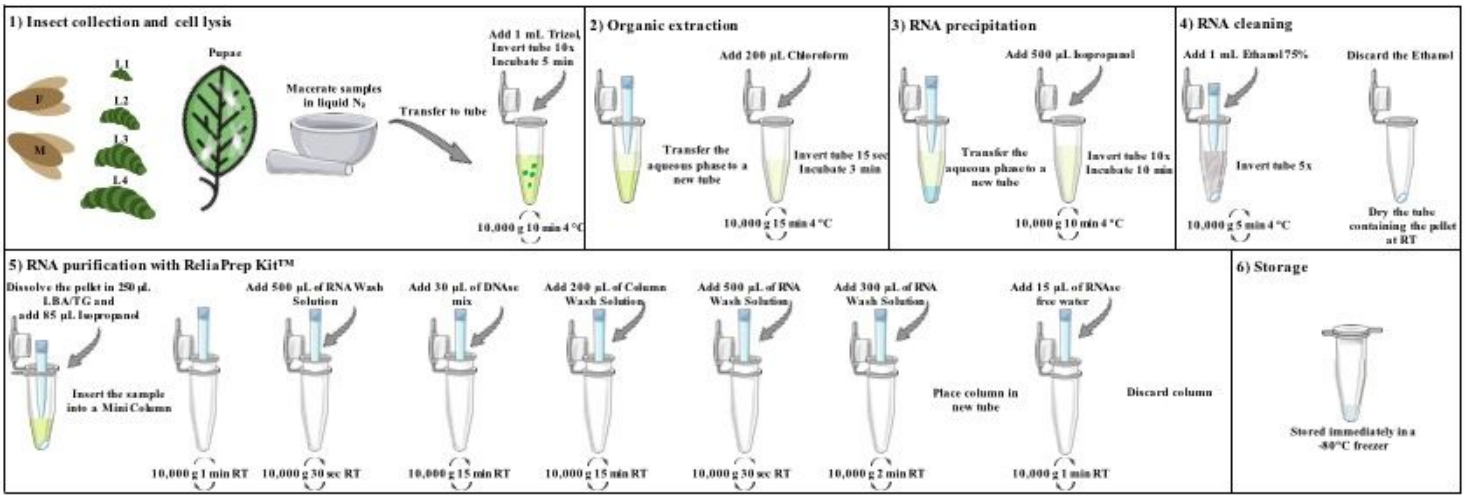

Figure 2

Synthesis of the steps of the optimized extraction of total RNAI: 1) Cell lysis; 2) Organic extraction; 3) RNA precipitation; 4) RNA cleaning; 5) RNA purification with ReliaPrep RNA Tissue Miniprep System; 6) Storage.

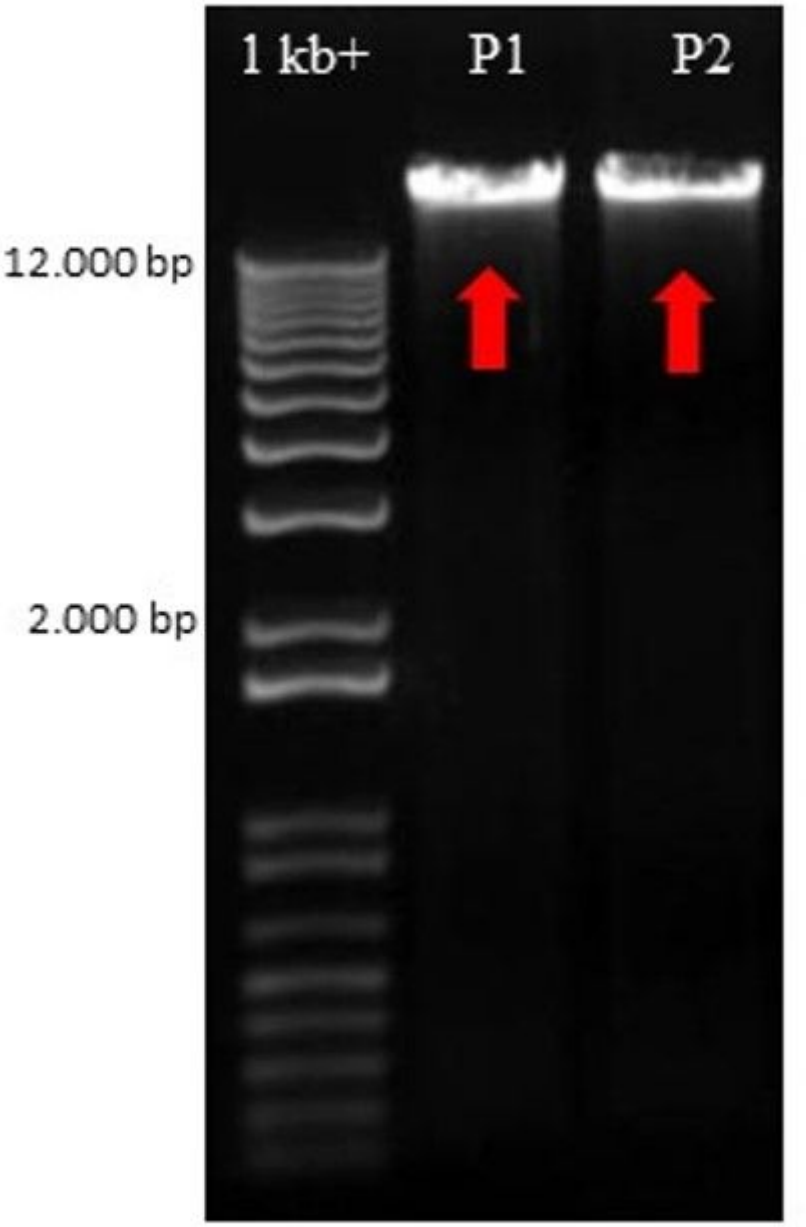

Figure 3

Electrophoresis in 0.8% agarose gel stained with ethidium bromide after 1 hour and 30 minutes of running at 50 volts (V). The marker used was the 1Kb Plus DNA Ladder (Invitrogen) (well 1). Samples P1 and P2 with gDNA bands larger than 12.000 bp (red arrows).

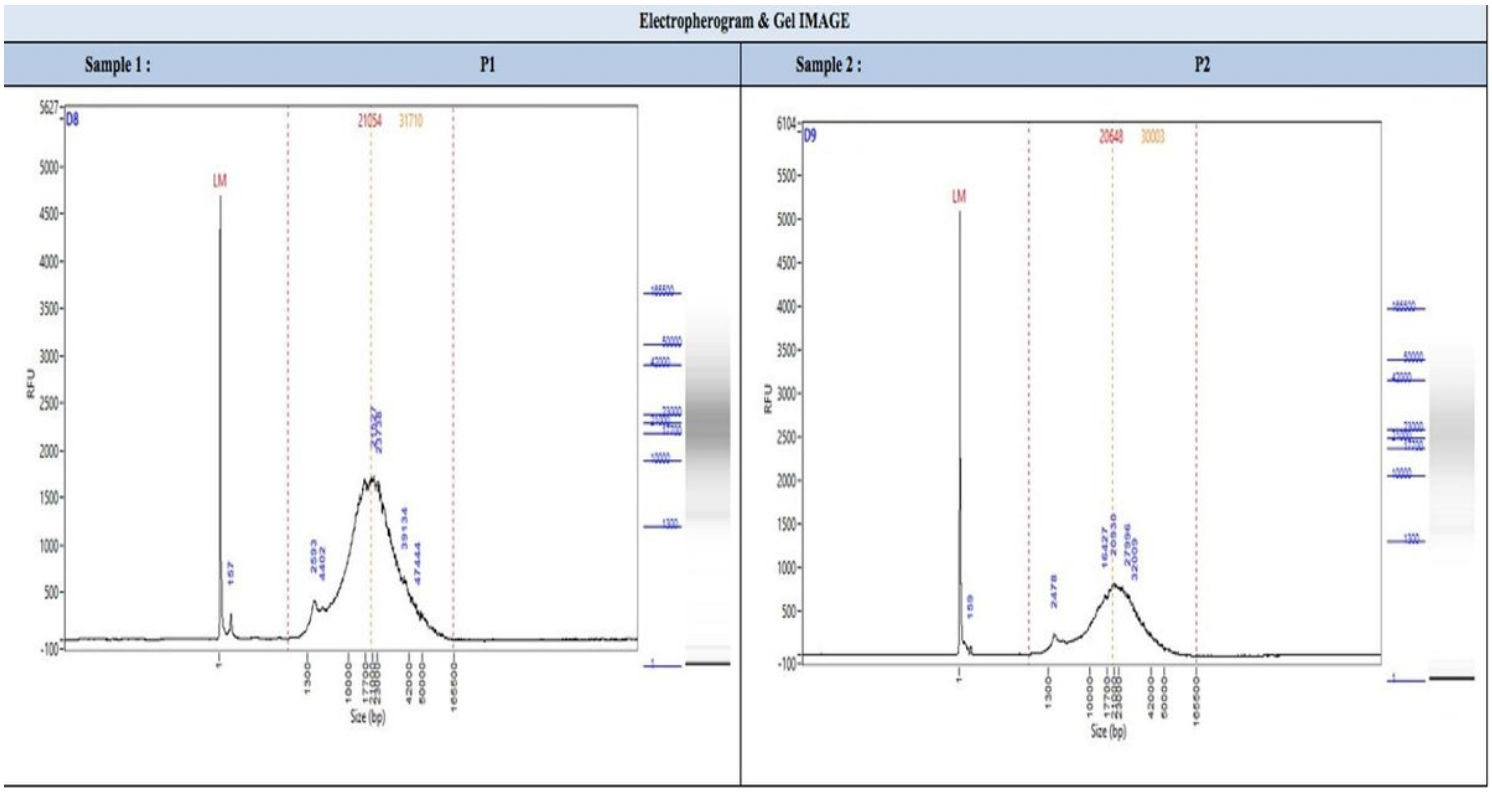

Figure 4

Electropherogram of the quality control of gDNA samples estimated using Femto Pulse (Agilent). P1: 46.3% of fragments above 20.000 bp; P2: 49.9% of fragments above 20.000 bp, ideal for PacBio sequencing.

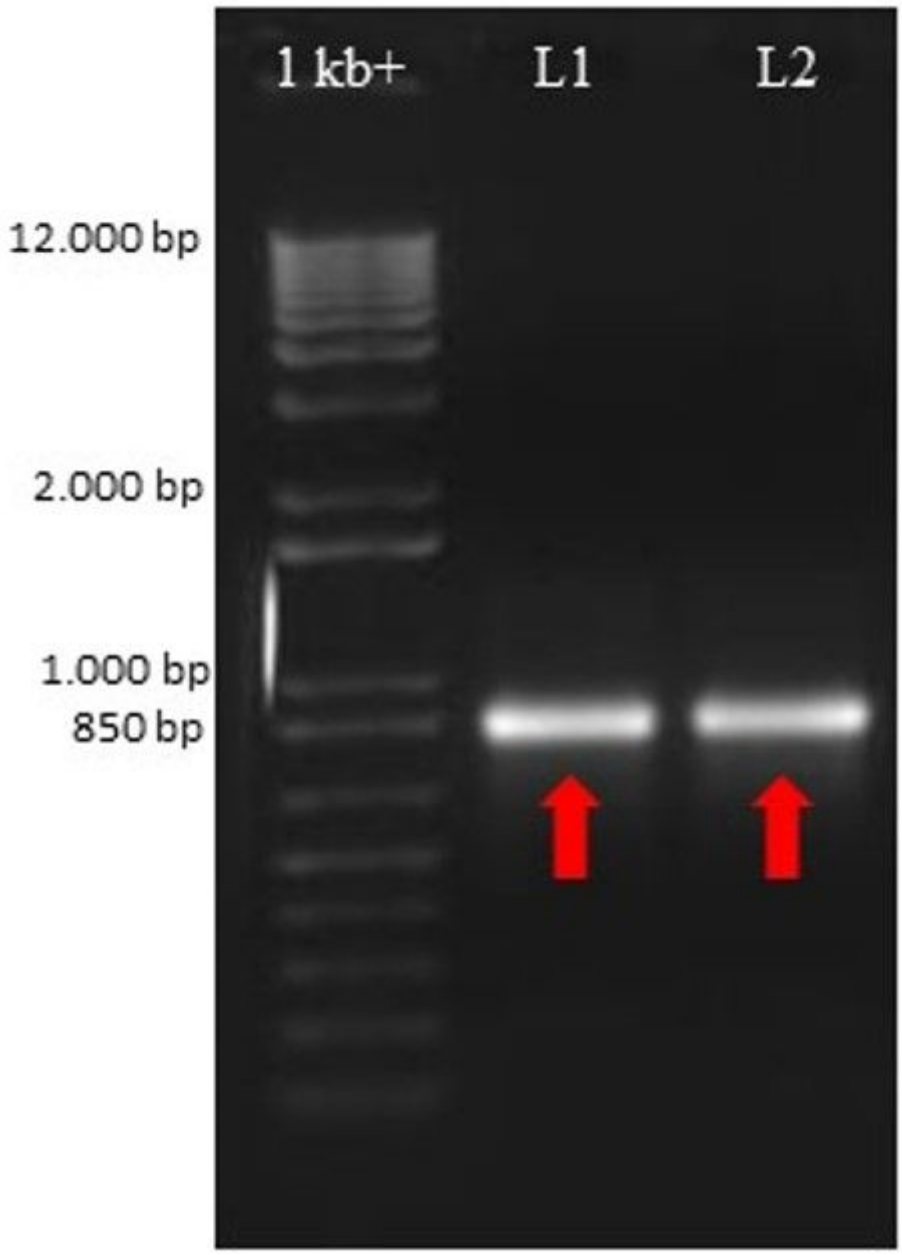

Figure 5

Electrophoresis in 1,2% non-denaturing agarose gel electrophoresis, stained with ethidium bromide after 2 hours of running at 50 volts (V). The marker used was the 1Kb Plus DNA Ladder (Invitrogen) (well 1). rRNA bands corresponding to the larval instars L1 and L2 with approximately 850 bp (red arrows), representing the remaining stages of development.

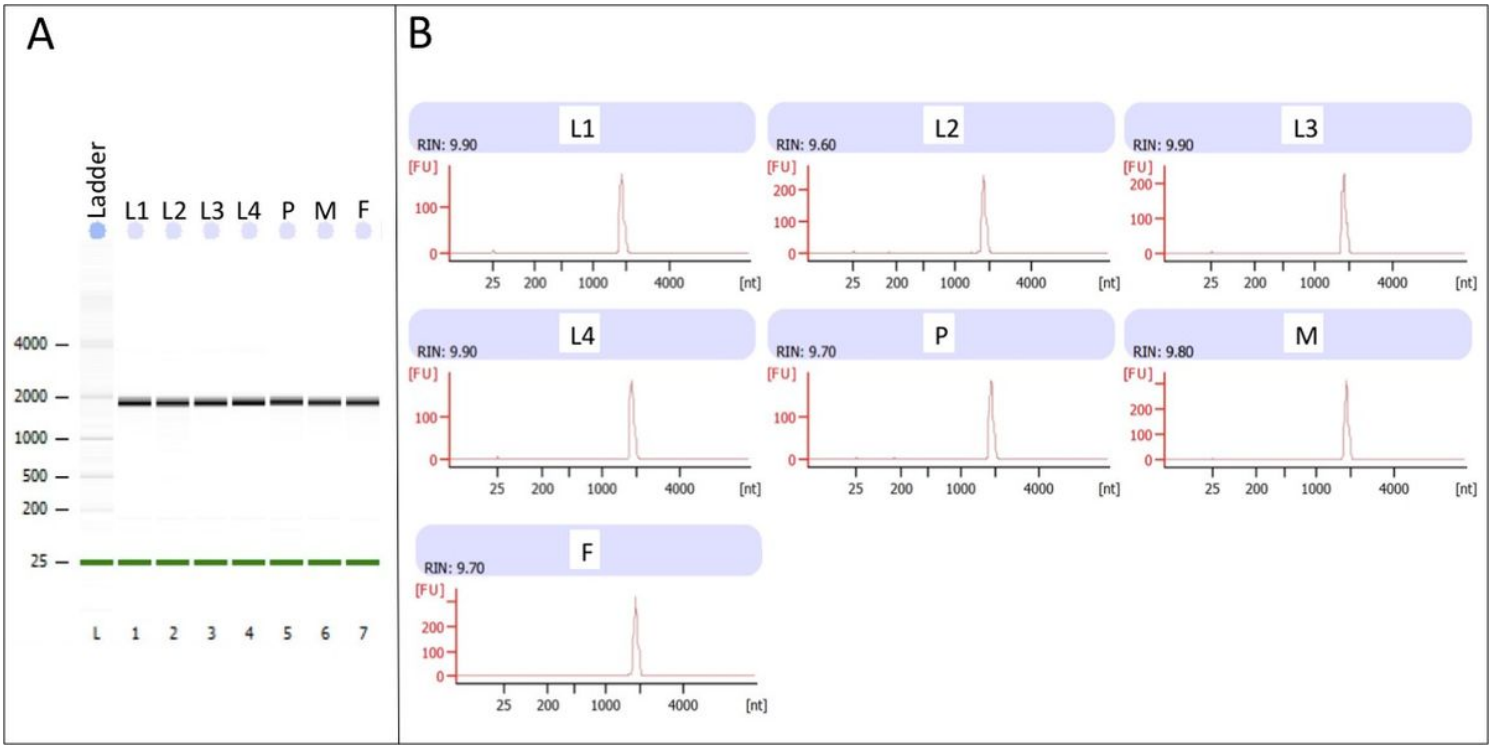

Figure 6

(A) Electropherogram generated by Bioanalyzer showing an overlap of the 28S and 18S subunits in a single band with 1.900 bp. Ladder corresponds to the standard RNA size of the picochip. (B) distribution peaks of RNA fragments with all RIN values

above 9.5 at all stages of development. L1 (larval instar 1), L2 (larval instar 2), L3 (larval instar 3), L4 (larval instar 4), P (pupae), M (males) and F (females).

Sample	A260/A280	A260/A230	Concentration (ng/µL)	Concentration (ng/µL)
			Nanodrop	Qubit
P1	1,83	2,22	360,8	309,6
P1	1,88	2,25	370,1	260,0

Figure 7

Table 1: Absorbance ratios and gDNA concentrations of *L. coffella* pupae quantified in Nanodrop and Qubit, respectively.

Sample	A260/A280	A260/A230	Concentration (ng/µL) Qubit
L1	2,18	2,26	876
L2	2,17	2,21	756
L3	2,17	2,27	816
L4	2,19	2,27	1159
P	2,15	2,32	542
M	2,19	2,28	732
F	2,20	2,32	634

Figure 8

Table 2: Absorbance ratios and total RNA concentration of the 7 stages of development from *L. coffella* quantified on Nanodrop and Qubit, respectively.