

Secuenciamiento del ADN genómico de *Myrciaria dubia* “camu-camu”

Juan Carlos Castro Gómez¹, J. Dylan Maddox^{2,3}, Marianela Cobos⁴, Sixto A. Imán⁵, Mario Pinedo⁶

¹Unidad Especializada de Biotecnología, Centro de Investigaciones de Recursos Naturales de la Amazonía (CIRNA), Universidad Nacional de la Amazonía Peruana (UNAP), Iquitos-Perú.

²Pritzker Laboratory for Molecular Systematics and Evolution, The Field Museum of Natural History, Chicago, IL, USA

³Environmental Sciences, American Public University System, Charles Town, WV, USA

⁴Laboratorio de Biotecnología y Bioinformática, Universidad Científica del Perú (UCP), Iquitos-Perú.

⁵Área de Conservación de Recursos Genéticos de Plantas, Instituto Nacional de Innovación Agraria (INIA), Iquitos-Perú.

⁶Instituto de Investigaciones de la Amazonía Peruana (IIAP)

RESUMEN

Myrciaria dubia “camu-camu”, es una especie nativa de la Amazonía caracterizada por sus frutos con alto contenido de ácido L-ascórbico (vitamina C) y otras sustancias bioactivas. Sin embargo, los estudios moleculares para esta especie aún son muy limitados. Pero, para ampliar nuestros conocimientos a nivel genómico, una etapa indispensable es la purificación del ADN genómico. Por tanto, los objetivos de este componente del Proyecto fueron: 1) purificar el ADN genómico a partir de hojas de *M. dubia*, 2) secuenciar y ensamblar el genoma cloroplástico de *M. dubia*. El ADN se purificó con un protocolo estandarizado en la Unidad Especializada de Biotecnología – CIRNA-UNAP, la calidad y cantidad del ADN se evaluó mediante análisis espectrofotométrico y electroforético. En base al ADN obtenido se elaboraron librerías y se procedió a secuenciarlas con la tecnología de PacBio. Los ratios de calidad promedio fueron apropiados ($A_{260}/A_{230} = 2.25 \pm 0.07$, $A_{260}/A_{280} = 1.81 \pm 0.01$), indicando ausencia de contaminantes como polisacáridos, polifenoles y proteínas, respectivamente. Asimismo, la alta calidad se corroboró mediante los resultados del análisis electroforético, el cual muestra integridad de las bandas de ADN y ausencia de ARN. Se ha generado más de $6,0 \times 10^9$ nucleótidos de data, y al ser ensamblados se generaron 9,092 contigs. En base a esta información se obtuvo el genoma cloroplástico ensamblado y anotado. En conclusión, con las aproximaciones moleculares y bioinformáticas se ha purificado ADN genómico de alta calidad y secuenciado, ensamblado y anotado el genoma cloroplástico de *M. dubia*.

Palabras claves: Ácidos nucleicos, frutal nativo, plataforma de PacBio, secuenciamiento genómico

INTRODUCCIÓN

Myrciaria dubia “camu-camu” es un frutal nativo de la Amazonía que produce frutos ricos en diversos compuestos nutraceuticos. Por ejemplo, presentan un alto contenido de ácido L-ascórbico (vitamina C), el cual puede superar los dos gramos por 100 g pulpa [1,2]. Adicionalmente, contiene antocianinas [3], 1-metil malato [4], fenoles [5], compuestos con propiedades antígenotóxicas [6] y neuroprotectivas [7]. Aunque reportes recientes basados en el análisis de la expresión de genes [8] y del transcriptoma [9] han revelado algunos componentes moleculares responsables de la biosíntesis de varios de estos compuestos, aún los conocimientos a nivel molecular son limitados. En ese sentido, para ampliar los conocimientos moleculares a nivel genómico, una etapa clave para realizar los análisis genómicos es la purificación del ADN genómico.

El proceso de purificación de ADN genómico debe caracterizarse por permitirnos obtener ADN de alta calidad y una cantidad suficiente para todos los análisis moleculares al que será sometido. La alta calidad está directamente relacionada con su integridad (ADN sin degradar) y su pureza. La pureza nos indica ausencia de proteínas, polifenoles, polisacáridos, ARN y otros compuestos contaminantes que interfieren con los posteriores análisis, es decir impidiendo que se realice una cuantificación exacta o mediante la inhibición de reacciones enzimáticas como de las endonucleasas de restricción y de las ADN polimerasas [10,11]. Finalmente, el protocolo de purificación también debe caracterizarse por ser rápido, simple, de bajo costo y no emplear sustancias químicas tóxicas [12].

En la actualidad están publicados múltiples protocolos para purificar el ADN genómico de las plantas [13–16]. Sin embargo, debido a la complejidad bioquímica que presentan las plantas y aún a nivel de los tejidos de una misma planta, los protocolos desarrollados no son aplicables a todas las especies. Aunque, previamente hemos desarrollado un protocolo para purificar el ADN a partir de hojas de *M. dubia* [17], en este informe mostramos los resultados obtenidos con el mismo protocolo pero con algunas modificaciones. Asimismo, para ampliar los conocimientos moleculares a nivel genómico se ha procedido a secuenciar y ensamblar el genoma cloroplástico de esta especie. En conjunto estos resultados, nos permitirán sentar las bases para el futuro mejoramiento genético de *M. dubia*, pero con el soporte de la información genómica.

MATERIALES Y MÉTODOS

Material botánico

Las muestras biológicas (hojas tiernas) de *M. dubia* fueron obtenidas de la Colección de germoplasma de *Myrciaria dubia* “camu-camu”, de la Estación Experimental “El Dorado” del Instituto Nacional de Innovación Agraria (INIA), ubicado a la altura del km 25½ de la carretera Iquitos-Nauta. Las muestras obtenidas fueron transportadas en condiciones de refrigeración a la Unidad Especializada de Biotecnología, Centro de Investigaciones de Recursos Naturales de la Amazonía (CIRNA), Universidad Nacional de la Amazonía Peruana (UNAP), para su almacenamiento y realizar la purificación del ADN genómico.

Purificación del ADN genómico

Se realizó en base a Castro et al. [17] modificado, que constó de las siguientes etapas: se pesó 150 mg de hojas tiernas y trituró en un mortero, añadiendo gradualmente 3 mL de tampón de extracción (Tris-HCl 300 mM pH 8.0, NaCl 2.0 M, EDTA 50 mM, CTAB 2%, PVP 1%, β - mercaptoetanol 5%). El triturado se homogenizó en el vortex por 30 seg e incubó a 75 °C por 30 min, mezclando por inversión cada 10 min. Luego se añadió 1.0 mL de cloroformo:alcohol isoamílico (24:1), se mezcló en el vortex por 30 seg y centrifugó a 23000 x g a 4 °C por 10 min. Esta etapa se repitió una vez más. Al sobrenadante se añadió igual volumen de isopropanol frío e incubó a -20 °C por 10 min. Posteriormente, se centrifugó a 10000 x g a 4 °C por 5 min, se descartó el sobrenadante y el ADN genómico se lavó dos veces con etanol al 70%. Después de centrifugar a 10000 x g a 4 °C por 5 min, el ADN precipitado se secó a 40 °C por 5 min. Luego el ADN se disolvió con 100 μ L de tampón TE (Tris- HCl 10 mM, EDTA 1 mM), se añadió 1 U de ARNasa A e incubó a 37°C por 30 min. La ARNasa A se eliminó mediante extracción con cloroformo:alcohol isoamílico (24:1), tal como se describió previamente.

La calidad y cantidad del ADN genómico purificado se determinó de acuerdo a Sambrook et al.[18], empleando espectrofotometría UV/Vis (con un espectrofotómetro Spectronic Genesys 6) y electroforesis en gel de agarosa al 1%. El registro fotográfico de los geles se hizo con el sistema BioDocAnalyze System (Biometra, Alemania).

Secuenciamiento y ensamblado del genoma

El genoma fue secuenciado usando la plataforma PacBio® RS II system [11] y las secuencias fueron ensambladas de novo en base a secuencias relacionadas de cloroplastos de especies de la familia Myrtaceae empleando el programa SPADES v3.6.1 (<http://bioinf.spbau.ru/spades>) y SOAPdenovo2 (<https://sourceforge.net/projects/soapdenovo2/>). Los contigs correspondientes a los cloroplastos

fueron identificados usando el algoritmo de mapeo y ensamblados en Geneious R11 (<https://www.geneious.com/>). El genoma fue anotado usando Dual Organellar GenoMe Annotator (DOGMA) (<https://dogma.cccb.utexas.edu/>). Finalmente, el mapa circular del genoma cloroplástico fue generado con el programa Organellar Genome DRAW (<https://chlorobox.mpimp-goim.mpg.de/OGDraw.html>).

Análisis de datos

La calidad del ADN genómico purificado se estimó mediante espectrofotometría con la siguiente ecuación: Ratios de calidad: A_{260}/A_{230} y A_{260}/A_{280}

Para estimar la concentración del ADN genómico purificado mediante espectrofotometría se empleó la ecuación: $\text{ADN } \mu\text{g/mL} = A_{m_{260}} \times 50 \mu\text{g/mL} \times \text{factor de dilución}$

Donde, $A_{m_{260}}$: absorbancia de la muestra de ADN a 260 nm

Todos los datos obtenidos fueron analizados mediante estadística descriptiva empleando el programa IBM SPSS Statistic v 21.

RESULTADOS Y DISCUSIÓN

Purificación de ADN genómico

El protocolo para purificar el ADN genómico a partir de las hojas de *M. dubia* fue idóneo. Porque los ratios de calidad promedio fueron apropiados ($A_{260}/A_{230} = 2.25 \pm 0.07$, $A_{260}/A_{280} = 1.81 \pm 0.01$) (Figura 1), indicando ausencia de contaminantes como polisacáridos, polifenoles y proteínas, respectivamente [13,14,18–20]. Ratios similares fueron obtenidos para *M. dubia* [17] y otras especies de plantas tales como *Arabidopsis thaliana* [21], *Mauritia flexuosa* [22], *Phytolacca dodecandra* [23] y *Echinaceae Purpurea* [24]. De acuerdo a Porebski [19], los ratios de absorbancia bajos (≤ 1.5) indican que el ADN genómico está contaminado con polifenoles y otros metabolitos secundarios. Por otra parte, de acuerdo a Castro et al. [17] los ratios mayores de 2.0 pueden sugerirnos que hubo degradación del ADN, porque a mayor fragmentación del ADN más bases nitrogenadas absorben la luz ultravioleta a 260 nm, incrementándose la absorción a esta longitud de onda. Además, este incremento causa una sobreestimación de la concentración de esta macromolécula.

Por otra parte, la alta calidad del ADN genómico se corroboró mediante los resultados del análisis electroforético, el cual muestra integridad de las bandas de ADN genómico y ausencia de ARN (Figura 1). Asimismo, la ausencia de polisacáridos se evidencia porque el ADN no es atrapado en los pocillos del gel. Este problema sucede cuando los polisacáridos co-precipitan con el ADN en la etapa de precipitación con alcoholes (etanol o isopropanol), formando complejos insolubles y si esto ocurre los geles muestran que un grupo de moléculas de ADN genómico quedan atrapados en los pocillos [17]. Finalmente, cabe señalar que el rendimiento promedio del protocolo fue de $350 \pm 50 \mu\text{g}$ de ADN genómico por gramo de hoja, lo cual representa una cantidad de ADN para múltiples ensayos moleculares.

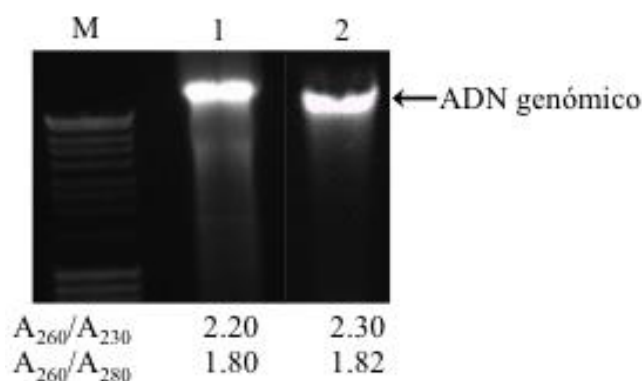


Figura 1. Resultados de los análisis electroforético y espectrofotométrico del ADN genómico purificado a partir de hojas de *M. dubia*. M: marcador de peso molecular, 1 y 2 muestras de ADN obtenidas por duplicado.

Secuenciamiento, ensamblado y anotación del genoma cloroplástico

Se ha secuenciado y pre-ensamblado el genoma de *Myrciaria dubia* con la tecnología de secuenciamiento de PacBio. En total se ha generado 6,602,587,428 nucleótidos de data, los que fueron ensamblados en un total de 9,092 contigs con una longitud máxima de 276,741 pares de bases. El N₅₀ de la longitud de los contigs fue de 36,556 (Figura 2).

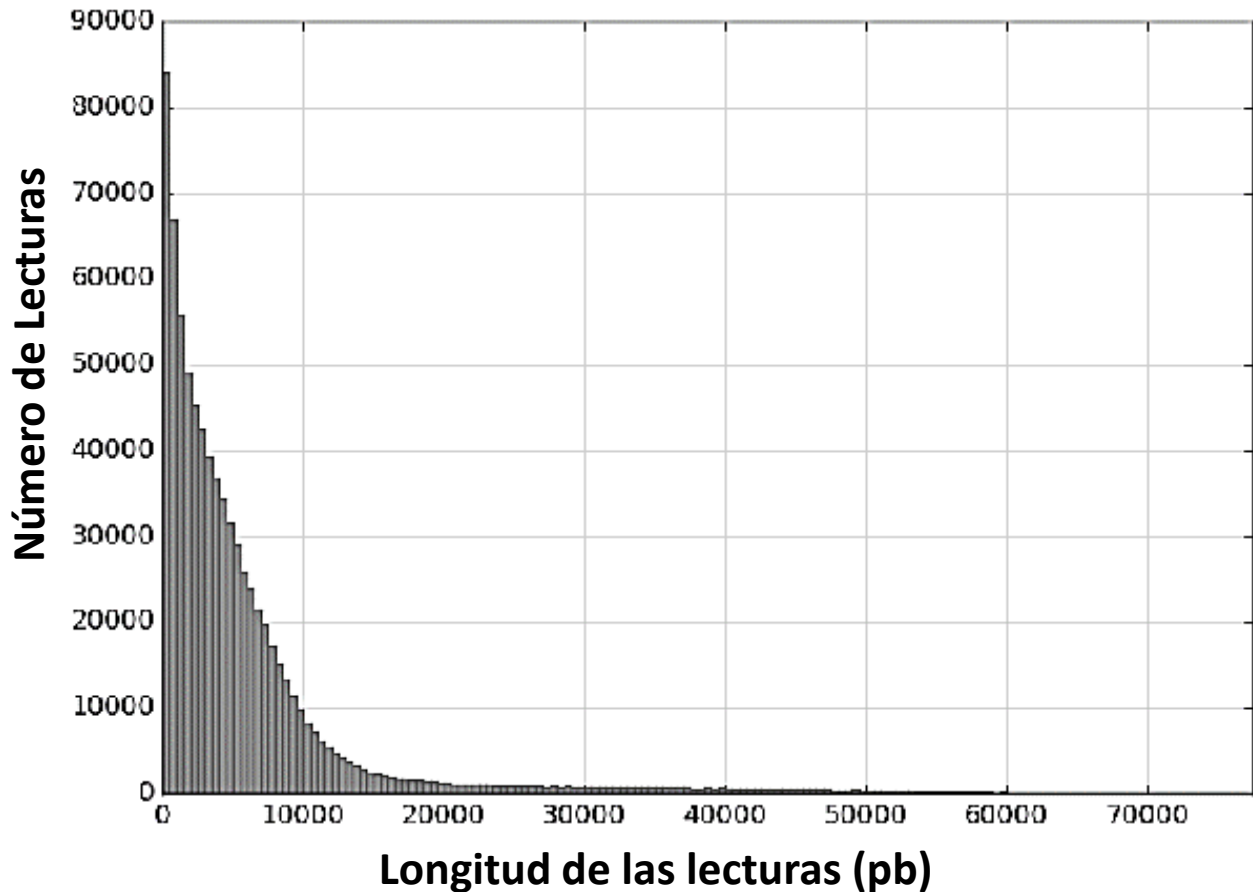


Figura 2. Distribución de las longitudes de las lecturas del ADN genómico secuenciado de *M. dubia* con tecnologías de secuenciamiento masivo.

Asimismo, luego de ensamblar y anotar el genoma cloroplástico de *M. dubia*, se aprecia que este circular de aproximadamente 160,000 pb de tamaño y está dividido en 4 regiones (LSC, SSC, IRA, IRB) y presenta genes involucrados en fotosíntesis, transcripción, traducción, etc. Lo cual es una característica típica de las Myrtaceae y de otras familias de plantas, que conservan esta estructura (Figura 3).

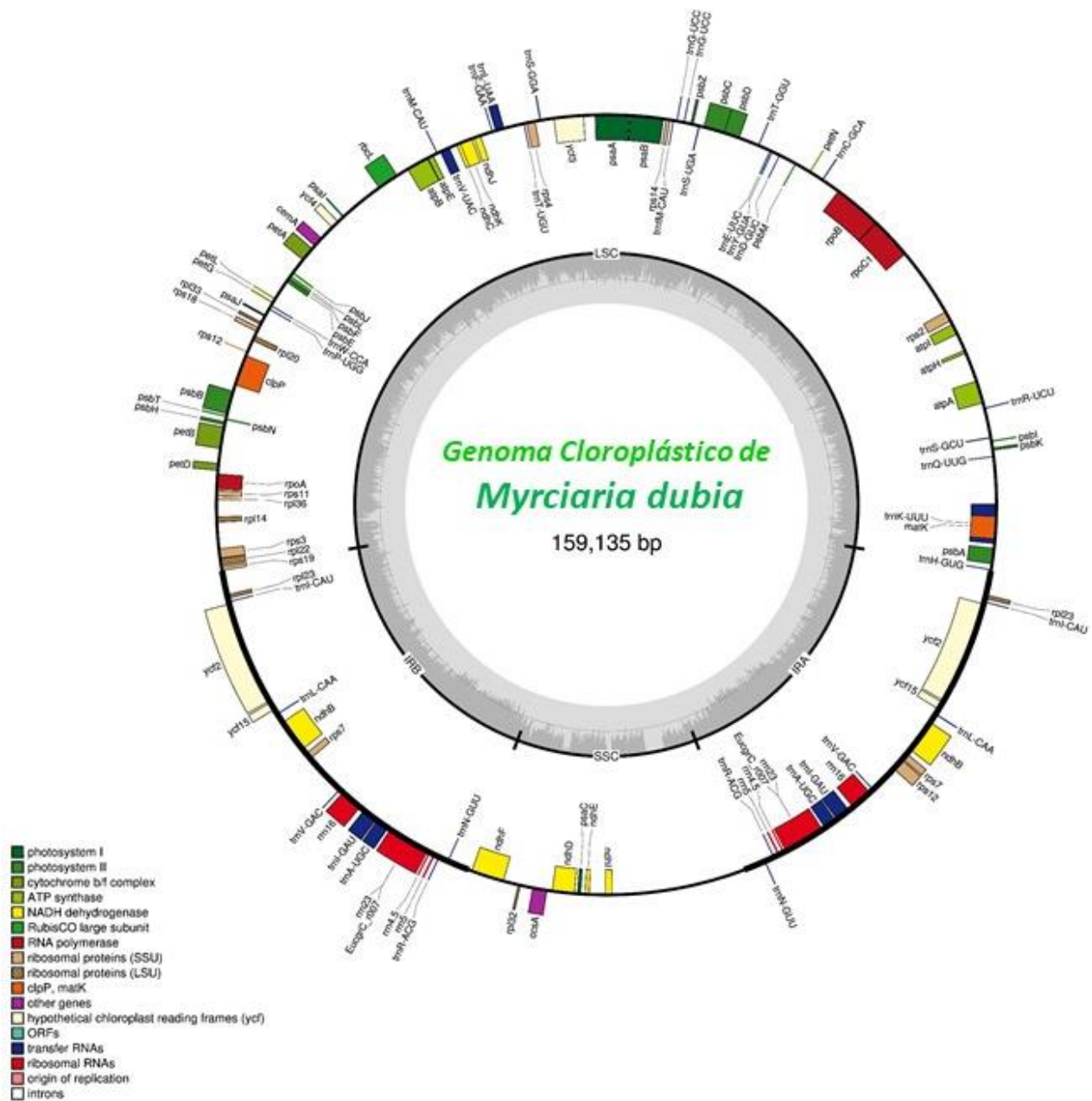


Figura 3. Mapa genético del Genoma Cloroplástico de *M. dubia*

CONCLUSIONES

Con el protocolo empleado se ha purificado el ADN genómico a partir de hojas de *M. dubia* “camu camu”, el cual se caracteriza por su alta calidad y cantidad apropiados para futuros análisis moleculares. Asimismo, se ha logrado secuenciar, ensamblar y anotar el genoma cloroplástico de *M. dubia*, que en un futuro inmediato dará soporte a procesos biotecnológicos cloroplásticos.

REFERENCIAS BIBLIOGRÁFICAS

1. Bradfield, RB, Roca, A. Camu-camu - a fruit high in ascorbic acid. J. Am. Diet. Assoc. 1964;44:28–30.
2. Imán S, Bravo L, Sotero V, Oliva C. Contenido de vitamina C en frutos de camu camu *Myrciaria dubia* (H.B.K) Mc Vaugh, en cuatro estados de maduración, procedentes de la Colección de Germoplasma del INIA Loreto, Perú. Sci. Agropecu. 2011;2:123–30.
3. Zanatta CF, Cuevas E, Bobbio FO, Winterhalter P, Mercadante AZ. Determination of anthocyanins from camu-camu (*Myrciaria dubia*) by HPLC-PDA, HPLC-MS, and NMR. J. Agric. Food Chem. 2005;53:9531–5.
4. Akachi T, Shiina Y, Kawaguchi T, Kawagishi H, Morita T, Sugiyama K. 1-methylmalate from camu-camu (*Myrciaria dubia*) suppressed D-galactosamine-induced liver injury in rats. Biosci. Biotechnol. Biochem. 2010;74:573–8.
5. Fracassetti D, Costa C, Moulay L, Tomás-Barberán FA. Ellagic acid derivatives, ellagitannins, proanthocyanidins and other phenolics, vitamin C and antioxidant capacity of two powder products from camu-camu fruit (*Myrciaria dubia*). Food Chem. 2013;139:578–88.
6. da Silva FC, Arruda A, Ledel A, Dauth C, Romão NF, Viana RN, et al. Antigenotoxic effect of acute, subacute and chronic treatments with Amazonian camu-camu (*Myrciaria dubia*) juice on mice blood cells. Food Chem. Toxicol. Int. J. Publ. Br. Ind. Biol. Res. Assoc. 2012;50:2275–81.
7. Azevêdo JCS, Borges KC, Genovese MI, Correia RTP, Vatted DA. Neuroprotective effects of dried camu-camu (*Myrciaria dubia* HBK McVaugh) residue in *C. elegans*. Food Res. Int. 2015;73:135–41.
8. Castro JC, Cobos M, Maddox JD, Imán SA, Egoavil A, Torres J, et al. Gene expression and enzyme activities of the D-mannose/L-galactose pathway influence L-ascorbic acid content in *Myrciaria dubia*. Biol. Plant. 2015;59:783–7.
9. Castro JC, Maddox JD, Cobos M, Requena D, Zimic M, Bombarely A, et al. De novo assembly and functional annotation of *Myrciaria dubia* fruit transcriptome reveals multiple metabolic pathways for L-ascorbic acid biosynthesis. BMC Genomics. 2015;16:997.
10. Pandey RN, Adams RP, Flournoy LE. Inhibition of random amplified polymorphic DNAs (RAPDs) by plant polysaccharides. Plant Mol. Biol. Report. 1996;14:17–22.
11. Koonjul PK, Brandt WF, Lindsey GG, Farrant JM. Inclusion of polyvinylpyrrolidone in the polymerase chain reaction reverses the inhibitory effects of polyphenolic contamination of RNA. Nucleic Acids Res. 1999;27:915–6.
12. Mogg RJ, Bond JM. A cheap, reliable and rapid method of extracting high-quality DNA from plants. Mol. Ecol. Notes. 2003;3:666–8.
13. Allen GC, Flores-Vergara MA, Krasynanski S, Kumar S, Thompson WF. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. Nat. Protoc. 2006;1:2320–5.
14. John ME. An efficient method for isolation of RNA and DNA from plants containing polyphenolics. Nucleic Acids Res. 1992;20:2381.

15. Jobses DV, Hurley DL, Thien LB. Plant DNA Isolation: A Method to Efficiently Remove Polyphenolics, Polysaccharides, and RNA. *Taxon*. 1995;44:379–86.
16. Michiels A, Van den Ende W, Tucker M, Van Riet L, Van Laere A. Extraction of high-quality genomic DNA from latex-containing plants. *Anal. Biochem*. 2003;315:85–9.
17. Castro JC, Cobos M, Ramírez R, Imán S. Aislamiento de ADN genómico de *Myrciaria dubia* (HBK) “camu-camu” apropiado para análisis moleculares. *Cienc. Amaz*. 2012;2:7–15.
18. Sambrook, J, Fritsch, EF, Maniatis, T. *Molecular Cloning: A Laboratory Manual*. Second. Cold Spring Harbor Laboratory Press; 1989.
19. Porebski S, Bailey LG, Baum BR. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol. Biol. Report*. 1997;15:8–15.
20. Doyle JJ, and Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull*. 1987;19:11–5.
21. Oñate-Sánchez L, Vicente-Carbajosa J. DNA-free RNA isolation protocols for *Arabidopsis thaliana*, including seeds and siliques. *BMC Res. Notes*. 2008;1:93.
22. Gómez C, Carlos J, Jiménez N, Andersson G, Gutiérrez C, Alexander L, et al. Purificación de ADN genómico a partir de hojas, raíces y neumatóforos de *Mauritia flexuosa* “aguaje.” *Rev. Soc. Quím. Perú*. 2013;79:209–17.
23. Seman K, Bjornstad A, Stedje B. Genetic diversity and differentiation in Ethiopian populations of *Phytolacca dodecandra* as revealed by AFLP and RAPD analyses. *Genet. Resour. Crop Evol*. 2003;50:649–61.
24. Vural HC, Dağeri A. Optimization of DNA isolation for RAPD-PCR analysis of selected (*Echinaceae Purpurea* L. Moench) medicinal plants of conservation concern from Turkey. *J. Med. Plants Res*. 2009;3:1084–7.