Plant Cell Biotechnology and Molecular Biology 17(7&8):415-427; 2016 ISSN: 0972-2025

ISOLATION AND MOLECULAR CLONING OF GENES FROM *Myrciaria dubia* "camu-camu" WITH POTENTIAL USE FOR BIOTECHNOLOGICAL PRODUCTION OF VITAMIN C

JUAN C. CASTRO*, MARIANELA COBOS, J. DYLAN MADDOX, SIXTO A. IMÁN AND JORGE L. MARAPARA

Unidad Especializada de Biotecnología, Centro de Investigaciones de Recursos Naturales de la Amazonía (CIRNA), Universidad Nacional de la Amazonía Peruana (UNAP), Pasaje Los Paujiles S/N, Iquitos, Postal Code: 16024, Perú [JCC, JLM].

Laboratorio de Biotecnología y Bioenergética, Universidad Científica del Perú (UCP), Av. Abelardo Quiñones km 2.5, Iquitos, Postal Code: 16024, Perú [MC].

The Field Museum of Natural History, Pritzker Laboratory for Molecular Systematics and Evolution, 1400 S. Lake Shore Drive, Chicago, IL, Postal Code: 60605, USA [JDM].

Environmental Sciences, American Public University System, Charles Town, WV, Postal Code: 25414, USA [JDM].

Instituto Nacional de Innovación Agraria (INIA), Estación Experimental San Roque, Área de Conservación de Recursos Fitogenéticos, Calle San Roque # 236, Iquitos, Postal Code: 16024, Perú [SAI].

[*For Correspondence: E-mail: juanccgomez@yahoo.es]

ABSTRACT

***Myrciaria dubia* "camu-camu" is a rich source of several bioactive phytochemicals and vitamin C (L-ascorbic acid, AsA). To gain insights about the genes involved in AsA biosynthesis in this plant species and consequently with potential use for its biotechnological production, here we report the isolation and molecular cloning of partial gene sequences of the D-mannose/L-galactose pathway. Degenerate primers designed by the multiple sequence alignment of related plant species were used to isolate in *M. dubia* the partial sequences of the six D-mannose/L-galactose pathway genes (*GMP, GME, GGP, GPP, GDH*, and *GLDH*). The deduced protein sequences of the six genes have more than 81% sequence identity to rosids and asterids species, with a closer phylogenetic relationship to *Eucalyptus grandis*. In conclusion, gene sequences of the D-mannose/L-galactose pathway involved in AsA biosynthesis of *M. dubia* were successfully isolated and cloned and the phylogenetic analysis indicated that these genes have been relatively well conserved throughout of plant evolution, reflecting the importance of the enzymes of this metabolic pathway for plant growth and survival. Additionally, the isolation and cloning of these genes allow us to implement systems for biotechnological production of AsA.**

Keywords: Gene cloning; molecular biotechnology; molecular phylogeny; L-ascorbic acid biosynthesis; tropical fruit.

INTRODUCTION

Myrciaria dubia (Kunth) McVaugh "camu camu" is an Amazonian plant that produces several bioactive compounds such as aldose reductase inhibitors (Ueda et al., 2004), anthocyanins (Zanatta et al., 2005), 1-methyl-malate (Akachi et al., 2010), ellagic acid derivatives and other phenolics (Fracassetti et al., 2013), but is economically important due to its high vitamin C (L-ascorbic acid, AsA) content in fruits (Bradfield and Roca, 1964), which can contain as much as 2 g of AsA per 100 g of pulp (Imán et al., 2011). Abundant variation in AsA content both among different tissue types of the same individual and between individuals (Castro et al., 2013a),

however, was largely responsible for the precipitous decline of *M. dubia* as an economically important export of the Peruvian Amazon because of the inability of local, small-scale farmers to cultivate products of uniform AsA content.

Various studies have shown that the high variation in AsA content in plant tissues is influenced by a variety of genetic and environmental factors (Davey et al., 2006; Roselló et al., 2011). These factors, directly or indirectly, through molecular mechanisms that are currently under investigation, influence the metabolic pathways of AsA biosynthesis (Zhang et al., 2009; Wang et al., 2013; Conklin et al., 2013). Although a combination of several experimental approaches (e.g., radiolabelling, mutant analysis, transgenic manipulation) provides evidence for multiple pathways of AsA biosynthesis in plants, the D-mannose/L-galactose pathway is the best characterized and commonly considered the most important (Wheeler et al., 1998; Valpuesta and Botella, 2004; Ishikawa et al., 2006).

M. dubia gene sequences of the D-mannose/L-galactose pathway are lacking, and thus preclude molecular studies to determine the influence of genetic factors on AsA content variation. Additionally, the isolation and cloning of these genes allow us to implement systems for biotechnological production of AsA. Our aim of this study, therefore, was to isolate and molecular cloning of partial sequences of the D-mannose/L-galactose pathway genes from *M. dubia*.

MATERIALS AND METHODS

Plant Material

Mature fruits were collected and subsequently stored at -80°C from the *M. dubia* germplasm bank (03°57'17"S, 73°24'55"W) at the Instituto Nacional de Innovación Agraria of Peru, department of Loreto. This germplasm bank, established approximately 20 years ago, consists of 43 representative accessions of genetic variability of *M. dubia* belonging to eight major river basins of the Loreto Region (Nanay, Itaya, Napo, Ucayali, Putumayo, Curaray, Tigre and Amazonas).

Total RNA Isolation and cDNA Synthesis

Total RNA was isolated from mature fruits (fruit pulp and peel) using the CTAB method, solvent extractions, and DNase treatment as described by Castro et al. (2013b). RNA quality and quantity were assessed by standard OD measurement and formaldehyde denaturing gel electrophoresis (Sambrook et al., 2012) (see Fig. S1 in Additional File 1). Single-stranded cDNA was obtained from 1.5 µg of total RNA using MuLV Reverse Transcriptase and oligo$(dT)_{16}$ following the manufacturer's instructions (Applied Biosystems). Reaction conditions were as follows: 25°C for 10 min, 42°C for 1 h, enzyme denaturing at 95°C for 5 min.

Degenerate Primer Design

In order to isolate partial gene sequences of the Smirnoff-Wheeler pathway (i.e., GDP-D-mannose pyrophosphorylase [GMP], GDP-D-mannose-3',5'-epimerase [GME], GDP-L-galactose phosphorylase [GGP], L-galactose-1-phosphate phosphatase [GPP], L-galactose dehydrogenase [GDH], and L-galactono-1,4-lactone dehydrogenase [GLDH]) full-length mRNA sequences from different plants species were obtained from the GenBank database of the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). These sequences were aligned with Clustal Omega software (Sievers et al., 2011) and degenerate primer pairs were designed on the basis of conserved coding regions using SC Primer software (Jabado et al., 2006). Primers were designed with the following specifications: amplicon size = 250-1000 bp, optimum primer length = 18 to 22 nt, primer Tm = 58-62°C, Tm difference ≤ 3°C, primer GC% = 45-55%, with all other parameters set at default (see Table S1 in Additional File 2).

Isolation of Partial Sequences of the D-mannose/L-galactose Pathway Genes

Partial gene sequences were amplified from a pool of cDNA templates (from fruits pulp and peel), with degenerate PCR primers using a MasterCycler ep Gradient (Eppendorf) thermal cycler. Each PCR reaction (20 µL) contained 1x PCR buffer, 3 mM $MgCl_2$, 0.2 mM of each dNTP,

0.5-1.0 μM degenerated primer pair, 0.4 U Taq polymerase (Fermentas), and 2 μL cDNA template. The PCR protocol consisted of an initial denaturing step for 5 min at 95ºC, followed by 40 cycles of 45 s at 95ºC, 30 s at the primer specific annealing temperature (i.e., 45-60ºC), 1 min at 72ºC, and a final extension step for 10 min at 72ºC. The amplified products were resolved on a 2% agarose gel, and products of the expected size were excised, gel purified using the Quick Gel Extraction and PCR Purification Combo Kit (Invitrogen, CA, USA), and ligated into a pCR®2.1-TOPO cloning vector (Invitrogen, CA, USA) according to the manufacturer's instructions. The ligation product was introduced into chemically competent *E. coli* via the heat-shock method. Gene sequences were confirmed by sequencing independent clones on both strands using M13 forward and reverse primers, and verifying their identity with BLAST analysis (Zhang et al., 1997).

Bioinformatics Analysis

Searches for similarity were performed in the GenBank using the BLASTX algorithm (Zhang et al., 1997). Multiple alignments of the deduced protein sequences were completed with Clustal Omega (Sievers et al., 2011), and phylogenetic trees were constructed with MEGA software 6.0 (Tamura et al., 2013) using the Neighbor-Joining method (Saitou and Nei, 1987) with distances derived from the Jones-Taylor-Thornton model (Jones et al., 1992) and 1,000 bootstrap replicates.

RESULTS AND DISCUSSION

Metabolic pathways for AsA biosynthesis in plants have been recently elucidated. Pioneering work by Isherwood et al. (1953) showed that germinating cress seedlings (*Lepidium sativum*) fed with metabolic precursors (i.e., L-galactono-1,4-lactone, L-gulono-1,4-lactone, D-galacturonic acid) significantly increased its AsA content. Subsequently, radioisotopic labeling experiments found that conversion of metabolic precursors to AsA in plants proceeded directly (Loewus, 1963). In 1998 the first biosynthetic pathway (i.e., D-mannose/L-galactose pathway) for AsA in plants was discovered (Wheeler et al., 1998). This pathway is the best characterized because it is supported by robust biochemical and molecular genetic evidence (Valpuesta and Botella, 2004; Cruz-Rus et al., 2011; Ren et al., 2013; Hancock et al., 2003; Huang et al., 2014). Moreover, other metabolic pathways were proposed for AsA biosynthesis in plants (Valpuesta and Botella, 2004).

Given the D-mannose/L-galactose pathway is typically considered the most important for AsA biosynthesis in plants (Wheeler et al., 1998; Smirnoff and Wheeler, 2000; Gallie, 2013), in this research using bioinformatics tools and molecular biology techniques we successfully amplified partial gene sequences of *M. dubia* from the six genes that comprise this metabolic pathway (Fig. 1). PCR products (266 to 929 bp) were isolated, cloned, sequenced, and deposited in GenBank with the following accession numbers: *GMP*: KJ502650.1, *GME*: KJ502651.1, *GGP*: KJ502652.1, *GPP*: KJ502653.1, *GDH*: KJ502654.1, and *GLDH*: KJ502655.1. Similar strategies were used to isolated genes of this metabolic pathway from *Brassica oleracea* (Østergaard et al., 1997), *Cucumis melo* (Pateraki et al., 2004), *Fragaria ananassa* (Do Nascimento et al., 2005), and *Malpighia glabra* (Badejo et al., 2009). In addition, the isolation of these gene sequences from cDNA of fruits suggests that this organ have capabilities for AsA biosynthesis in *M. dubia*, but has yet to be experimentally demonstrated. These results are in agreement with studies in other plant species that have demonstrated that the fruits, and other organs and tissues are also capable of AsA biosynthesis (Ren et al., 2013; Badejo et al., 2009; Xu et al., 2013; Hancock et al., 2007).

Homology search of these sequences in the All non-redundant GenBank CDS translations, PDB, SwissProt, PIR and PRF databases shown that the gene sequences isolated from *M. dubia* have homology with the D-mannose/L-galactose pathway gene of several plant species, and by comparing with the corresponding gene sequences of *Arabidopsis thaliana* is evident that homology exist (Table 1). Results showing low E values ($1x10^{-76}$ - $5x10^{-49}$) and high identity values (82-94%) between sequences pairs demonstrate that the isolated cDNA segments from tissues of

M. dubia are similar to the D-mannose/L-galactose pathway gene sequences.

Due to the gene sequences isolated having a key role in AsA biosynthesis further work is required to obtain the complete coding sequences. This goal can be achieved with the gold standard technique rapid amplification of cDNA ends (Frohman et al., 1988) at which point, experiments to express these genes in heterologous systems (i.e., *Escherichia coli*, *Chlamydomonas reinhardtii*, etc) for meticulous biochemical and structural analysis of the enzymes encoded will be possible. Finally, the definition of this biosynthetic pathway in *M. dubia* should allow engineering for increased AsA production, thus increasing their nutritional value and stress tolerance. Also, these genes could be used for the biotechnological production of AsA.

To further characterize the D-mannose/L-galactose pathway gene sequences isolated from *M. dubia*, we conducted a multiple alignment (see Figs. S2-S7 in Additional file 3) and phylogenetic analysis to compare the deduced protein sequences of the enzymes from rosids and asterids (Fig. 2). The sequences showed a high degree of conservation among all the species, but the deduced protein sequences of *M. dubia* were most similar to *Eucalyptus grandis* (i.e., GMP, GGP, GPP, GDH and GLDH) and *Malpighia glabra* (GME). This is due to that these species belong to the Rosid clade, which also has a strong support on genome-scale phylogenetic studies (Jansen et al., 2007; Wang et al., 2000). In addition, is evident that the phylogenetic analysis among species based on GME sequences showed a closed relationship in comparison with the other genes. This is attributed to that this gene, coding a nucleotide-diphospho-sugar interconversion enzyme, was originated in early plant evolution from a progenitor gene of an ancient prokaryote (Yin et al., 2011). In contrast, the genes specific for L-ascorbic acid biosynthesis were obtained more recently in plant evolution (Wheeler et al., 2015). The high conservation of the D-mannose/L-galactose pathway genes among plant species is likely because AsA is a widespread and efficient antioxidant that has multiple functions in plants (Smirnoff and Wheeler, 2000; Gallie, 2013; Smirnoff, 2000). AsA has a small cost in terms of synthesis and toxicity, and its benefits include defence of the glutathione pool and appropriate working of a variety of biocatalysts, and thus it has been hypothesized that AsA was subject to selection pressures during evolution (Gest et al., 2013).

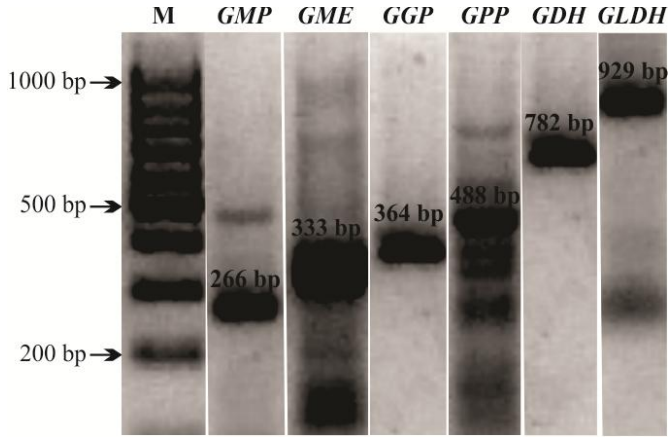

Fig. 1. Gel electrophoresis of amplicons obtained with degenerate primers of the Smirnoff-Wheeler pathway genes from *M. dubia*. M: molecular size marker 100-1000 bp, *GMP*: GDP-D-mannose pyrophosphorylase, *GME*: GDP-D-mannose-3',5'-epimerase, *GGP*: GDP-L-galactose phosphorylase, *GPP*: L-galactose-1-phosphate phosphatase, *GDH*: L-galactose dehydrogenase, and *GLDH*: L-galactono-1,4-lactone dehydrogenase

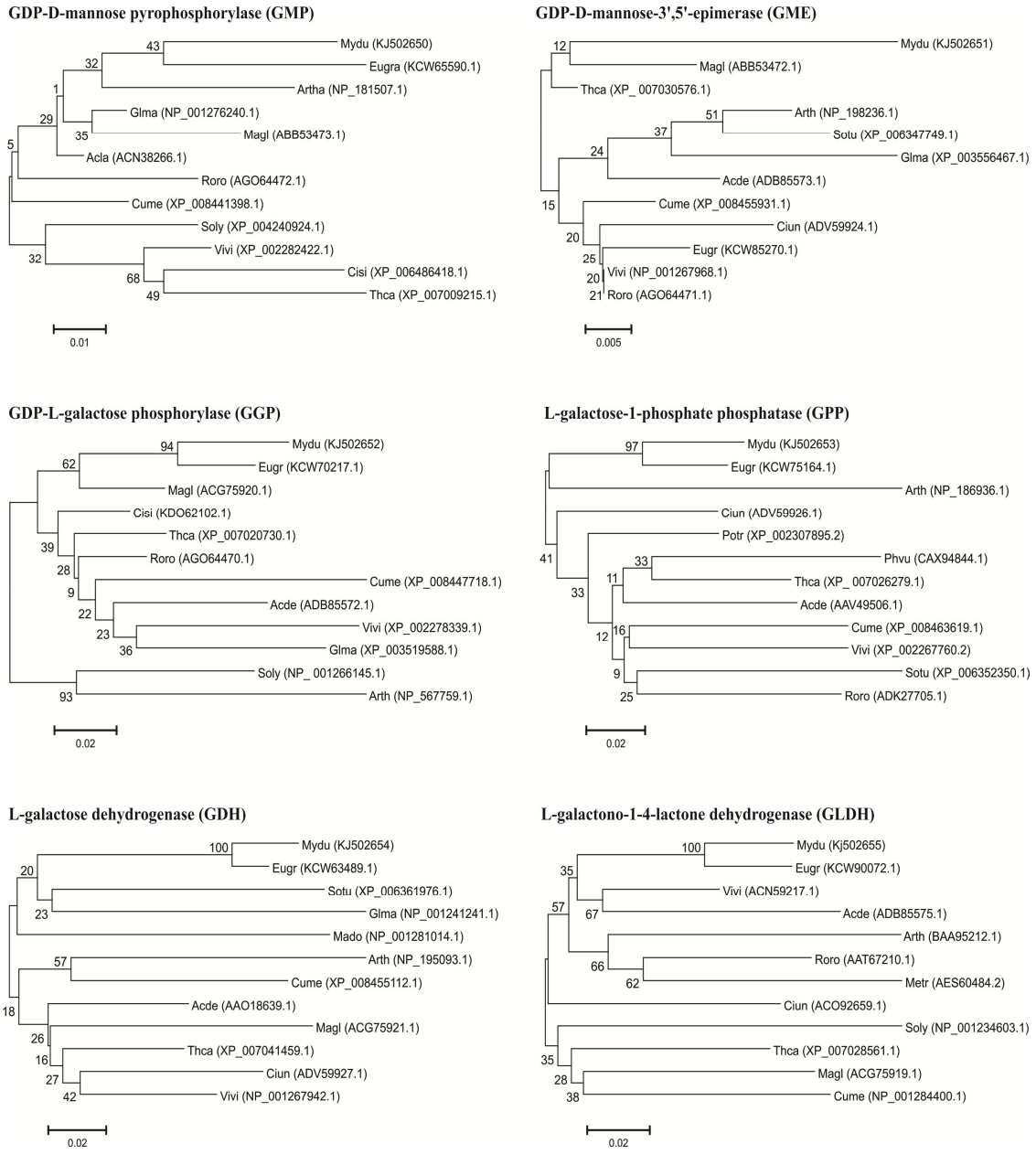

Fig. 2. Phylogenetic trees showing the relationship of the deduced protein sequences of the Smirnoff-Wheeler pathway genes of *M. dubia* to homologous proteins from rosids and asterids. The distances are the proportion of amino acid substitutions and the bootstrap values based on 1000 samples are shown. Acde: *Actinidia deliciosa*, Arth: *Arabidopsis thaliana*, Cisi: *Citrus sinensis*, Ciun: *Citrus unshiu*, Cume: *Cucumis melo*, Eugr: *Eucalyptus grandis*, Glma: *Glycine max*, Magl: *Malpighia glabra*, Mado: *Malus domestica*, Metr: *Medicago truncatula*. Mydu: *Myrciaria dubia*, Phvu: *Phaseolus vulgaris*, Potr: *Populus trichocarpa*, Roro: *Rosa roxburghii*, Soly: *Solanum lycopersicum*, Sotu: *Solanum tuberosum*, Thca: *Theobroma cacao*, Vivi: *Vitis vinifera*

Table 1. Comparison of the D-mannose/L-galactose pathway gene sequences of *M. dubia* and *Arabidopsis thaliana* after homology analysis with Blastx

Description	Total score	Query cover (%)	E value	Identity (%)	Accession
GDP-D-mannose pyrophosphorylase (*GMP*)	169	99	$5x10^{-49}$	93	NP_181507.1
GDP-D-mannose-3',5'-epimerase (*GME*)	220	99	$4x10^{-68}$	94	NP_198236.1
GDP-L-galactose phosphorylase (*GGP*)	209	98	$8x10^{-63}$	83	NP_567759.1
L-galactose-1-phosphate phosphatase (*GPP*)	284	99	$2x10^{-93}$	83	NP_186936.1
L-galactose dehydrogenase (*GDH*)	434	98	$6x10^{-150}$	82	NP_195093.1
L-Galactono-1,4-lactone dehydrogenase (*GLDH*)	514	99	$1x10^{-176}$	82	BAA95212.1

CONCLUSION

Gene sequences of the D-mannose/L-galactose pathway involved in AsA biosynthesis of *M. dubia* were successfully isolated and cloned and the phylogenetic analysis indicated that these genes have been relatively well conserved throughout of plant evolution, reflecting the importance of the enzymes of this metabolic pathway for plant growth and survival. Additionally, the isolation and cloning of these genes allow us to implement systems for biotechnological production of AsA.

ACKNOWLEDGEMENTS

This work was supported by grants from CONCYTEC (Contract N°280-2010) and UNAP (R.R. 2011).

COMPETING INTERESTS

The authors have declared that no competing interests exist.

REFERENCES

Akachi, T., Shiina, Y., Kawaguchi, T., Kawagishi, H., Morita, T., Sugiyama, K. 2010. 1-methylmalate from camu-camu (*Myrciaria dubia*) suppressed D-galactosamine-induced liver injury in rats. Biosci. Biotechnol. Biochem. **74**:573-578.

Badejo, A.A., Fujikawa, Y., Esaka, M. 2009. Gene expression of ascorbic acid biosynthesis related enzymes of the Smirnoff-Wheeler pathway in acerola (*Malpighia glabra*). J. Plant Physiol. **166**:652-660.

Bradfield, R.B., Roca, A. 1964. Camu-camu - a fruit high in ascorbic acid. J Am Diet Assoc. **44**:28-30.

Castro, J.C., Gutierrez, F., Acuña, C., Cerdeira, L.A., Tapullima, A., Cobos, M., Imán, S.A. 2013a. Variación del contenido de vitamina C y antocianinas en *Myrciaria dubia* «camu camu». Rev. Soc. Quím. Perú. **79**:319-330.

Castro, J.C., Egoávil, A.C., Torres, J., Ramírez, R., Cobos, M., Imán, S.A. 2013b. Isolation of high-quality total RNA from leaves of *Myrciaria dubia* "camu-camu". Prep. Biochem. Biotechnol. **43**:527-538.

Conklin, P.L., DePaolo, D., Wintle, B., Schatz, C., Buckenmeyer, G. 2013. Identification of Arabidopsis VTC3 as a putative and unique dual function protein kinase: Protein phosphatase involved in the regulation of the ascorbic acid pool in plants. J. Exp. Bot. **64**:2793-2804.

Cruz-Rus, E., Amaya, I., Sánchez-Sevilla, J.F., Botella, M.A., Valpuesta V. 2011. Regulation of L-ascorbic acid content in strawberry fruits. J. Exp. Bot. **62**:4191-4201.

Davey, M.W., Kenis, K., Keulemans, J. 2006. Genetic control of fruit vitamin C contents. Plant Physiol. **142**:343-351.

Do Nascimento, J.R., Higuchi, B.K., Gómez, M.L., Oshiro, R.A., Lajolo, F.M. 2005. L-Ascorbate biosynthesis in strawberries: L-galactono-1,4-lactone dehydrogenase expression during fruit development and ripening. Postharvest Biol. Technol. **38**:34-42.

Fracassetti, D., Costa, C., Moulay, L., Tomás-Barberán, F.A. 2013. 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. **139**:578-588.

Frohman, M.A., Dush, M.K., Martin, G.R. 1988. Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA. **85**:8998-9002.

Gallie, D.R. 2013. L-ascorbic Acid: A multifunctional molecule supporting plant growth and development. Scientifica (Cairo). **2013**:795964.

Gest, N., Gautier, H., Stevens, R. 2013. Ascorbate as seen through plant evolution: The rise of a successful molecule? J. Exp. Bot. **64**:33-53.

Hancock, R.D., McRae, D., Haupt, S., Viola, R. 2003. Synthesis of L-ascorbic acid in the phloem. BMC Plant Biol. **3**:7.

Hancock, R.D., Walker, P.G., Pont, S.D., Marquis, N., Vivera, S., Gordon, S.L., Brennan, R.M., Viola, R. 2007. L-Ascorbic acid accumulation in fruit of *Ribes nigrum* occurs by in situ biosynthesis via the L-galactose pathway. Funct Plant Biol. **34**:1080-1091.

Huang, M., Xu, Q., Deng, X.X. 2014. L-Ascorbic acid metabolism during fruit development in an ascorbate-rich fruit crop chestnut rose (*Rosa roxburghii* Tratt). J. Plant Physiol. **171**:1205-1216.

Imán, S., Bravo, L., Sotero, V., Oliva, C. 2011. 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. **2**:123-130.

Isherwood, F., Chen, Y., Mapson, L. 1953. Synthesis of L-ascorbic acid in plants and animals. Nature. **171**:348-349.

Ishikawa, T., Dowdle, J., Smirnoff, N. 2006. Progress in manipulating ascorbic acid biosynthesis and accumulation in plants. Physiol. Plant. **126**:343-355.

Jabado, O.J., Palacios, G., Kapoor, V., Hui, J., Renwick, N., Zhai, J., Briese, T., Lipkin W.I. 2006. Greene SCPrimer: A rapid comprehensive tool for designing degenerate primers from multiple sequence alignments. Nucleic Acids Res. **34**:6605-6611.

Jansen, R.K., Cai, Z., Raubeson, L.A., Daniell, H., dePamphilis, C.W., Leebens-Mack, J., Müller, K.F., Guisinger-Bellian, M., Haberle, R.C., Hansen, A.K., Chumley, T.W., Lee, S.B., Peery, R., McNeal, J.R., Kuehl, J.V., Boore, J.L. 2007. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. Proc Natl Acad Sci. **104**:19369–19374.

Jones, D.T., Taylor, W.R., Thornton, J.M. 1992. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. **8**:275-282.

Loewus, F.A. 1963. Tracer studies on ascorbic acid formation in plants. Phytochemistry. **2**:109-128.

Østergaard, J., Persiau, G., Davey, M.W., Bauw, G., Montagu, M.V. 1997. Isolation of a cDNA coding for L-galactono-γ-lactone dehydrogenase, an enzyme involved in the biosynthesis of ascorbic acid in plants. Purification, characterization, cDNA cloning, and expression in yeast. J. Biol. Chem. **272**:30009-30016.

Pateraki, I., Sanmartin, M., Kalamaki, M.S., Gerasopoulos, D., Kanellis, A.K. 2004. Molecular characterization and expression studies during melon fruit development and ripening of L-galactono-1,4-lactone dehydrogenase. J. Exp. Bot. **55**:1623-1633.

Ren, J., Chen, Z., Duan, W., Song, X., Liu, T., Wang, J., Hou, X., Li, Y. 2013. Comparison of ascorbic acid biosynthesis in different tissues of three non-heading Chinese cabbage cultivars. Plant Physiol. Biochem. **73**:229-236.

Roselló, S., Adalid, A.M., Cebolla-Cornejo, J., Nuez, F. 2011. Evaluation of the genotype, environment and their interaction on carotenoid and ascorbic acid accumulation in tomato germplasm. J. Sci. Food Agric. **91**:1014-1021.

Saitou, N., Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406-425.

Sambrook, J., Green, M.R. 2012. Molecular Cloning: A Laboratory Manual. Fourth Edition. Cold Spring Harbor Laboratory Press; 2028.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thpmpson, J.D., Higgins, D.G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. **7**:539.

Smirnoff, N., Wheeler, G.L. 2000. Ascorbic acid in plants: Biosynthesis and function. Crit. Rev. Biochem. Mol. Biol. **35**:291-314.

Smirnoff, N. 2000. Ascorbic acid: Metabolism and functions of a multi-facetted molecule. Curr. Opin. Plant Biol. **3**:229-235.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. **30**:2725-2729.

Ueda, H., Kuroiwa, E., Tachibana, Y., Kawanishi, K., Ayala, F., Moriyasu, M. 2004. Aldose reductase inhibitors from the leaves of *Myrciaria dubia* (H. B. & K.) McVaugh. Phytomedicine. **11**:652-656.

Valpuesta, V., Botella, M.A. 2004. Biosynthesis of L-ascorbic acid in plants: New pathways for an old antioxidant. Trends Plant Sci. **9**:573-577.

Wang, H., Moore, M.J., Soltis, P.S., Bell, C.D., Brockington, S.F., Alexandre, R., Davis, C.C., Latvis, M., Manchester, S.R., Soltis, D.E. 2009. Rosid radiation and the rapid rise of angiosperm-dominated forests. Proc Natl Acad Sci. **106**:3853–3858.

Wang, J., Yu, Y., Zhang, Z., Quan, R., Zhang, H., Ma, L., Deng, X.W., Huang, R. 2013. Arabidopsis CSN5B interacts with VTC1 and modulates ascorbic acid synthesis. Plant Cell. **25**:625-636.

Wheeler, G.L., Ishikawa, T., Pornsaksit, V., Smirnoff, N. 2015. Evolution of alternative biosynthetic pathways for vitamin C following plastid acquisition in photosynthetic eukaryotes. eLife. **4**:e06369.

Wheeler, G.L, Jones, M.A., Smirnoff, N. 1998. The biosynthetic pathway of vitamin C in higher plants. Nature. **393**:365-369.

Xu, Y., Zhu, X., Chen, Y., Gong, Y., Liu, L. 2013. Expression profiling of genes involved in ascorbate biosynthesis and recycling during fleshy root development in radish. Plant Physiol. Biochem. **70**:269-277.

Yin, Y., Huang, J., Gu, X., Bar-Peled, M., Xu, Y. 2011. Evolution of plant nucleotide-sugar interconversion enzymes. Plos One. **6**:e27995.

Zanatta, C.F., Cuevas, E., Bobbio, F.O., Winterhalter, P., Mercadante, A.Z. 2005. Determination of anthocyanins from camu-camu (*Myrciaria dubia*) by HPLC-PDA, HPLC-MS, and NMR. J. Agric. Food Chem. **53**:9531-9535.

Zhang, J., Madden, T.L. 1997. PowerBLAST: A new network BLAST application for interactive or automated sequence analysis and annotation. Genome Res. **7**:649-656.

Zhang, W., Lorence, A., Gruszewski, H.A., Chevone, B.I., Nessler, C.L. 2009. AMR1, an Arabidopsis gene that coordinately and negatively regulates the mannose/l-galactose ascorbic acid biosynthetic pathway. Plant Physiol. **150**:942-950.

APPENDIX

ADDITIONAL FILES

Additional file 1

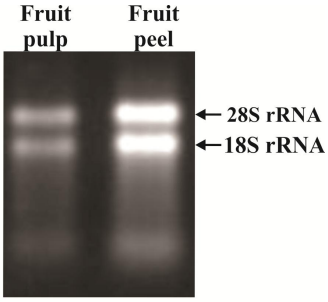

Fig. S1. Results of electrophoretic and spectrophotometric analysis of total RNA isolated from mature fruits of *M. dubia*. OD ratios (A_{260}/A_{230} and A_{260}/A_{280}) are shown as the mean ± SD (n = 3)

Additional file 2

Table S1. Lists of degenerate primers designed for amplification of the D-mannose/L-galactose pathway genes from *M. dubia*

Gene	GenBank accession number of mRNA used for multiple sequence alignment	Degenerate primer sequence	Product size (bp)
GDP-D-mannose pyrophosphorylase (GMP)	NM_001051330, EU180071.1, XM_002282386.1, DQ226694.1, XM_002524181.1,	GMPf 5'-AAYTAYCARCCDGAGGTKATG-3' GMPr 5'-GGYTCRTCMACCTTGGTHACCAT-3'	266
GDP-D-mannose-3',5'-epimerase (GME)	EU683446.1, HQ224948.1, HM998753.1, EF379384.1, FJ752238.1, GU339036.1,	GMEf 5'-GGCCHGCWGAGCCKCAAGAY-3' GMEr 5'-TGCTVACCATYTCRTCRCTTCC-3'	333
GDP-L-galactose phosphorylase (GGP)	XM_002316719.1, GU339036.1, XM_002530313.1, EF379384.1, EU683446.1.	GGPf 5'-TNAAYGARGGBCGHCAYYTKAAGAA-3' GGPr 5'-GCDCCHARRCTRTTRTAVYC-3'	364
L-galactose-1-phosphate phosphatase (GPP)	AB457584.1, HM234683.1, XM_002267724.2, FJ752240.1, AC154901.1	GPPf 5'-GGATTTGGTCACGGAAACT-3' GPPr 5'-CAATTCCACAGAGGTTCAGT-3'	488
L-galactose dehydrogenase (GDH)	319739584, 307136443, 146432258, 225382604, 170181407	GDHf 5'-CTTTGACACCTCCCCGTACT-3' GDHr 5'-GGCCATGACTGATTCTTCAC-3'	782
L-galactono-1,4-lactone dehydrogenase (GLDH)	AF252339.2, HM587128.1, AY547352.2, HM587129.1, EU683445.1	GLDHf 5'-GGTCACGCCTGCCAAGGGTA-3' GLDHr 5'-GGAACCTCAATCTTAGCCCAATGT-3'	929

Additional file 3

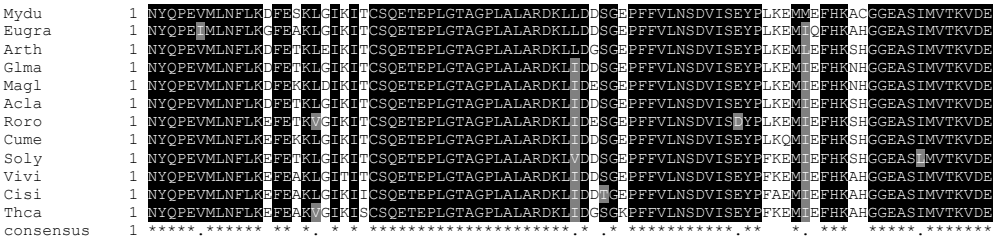

Fig. S2. Amino acid sequence alignment of GDP-mannose pyrophosphorylase (GMP) from *M. dubia* with sequences from other plant species. Mydu: *Myrciaria dubia* (KJ502650), Eugr: *Eucalyptus grandis* (KCW65590.1), Arth: *Arabidopsis thaliana* (NP_181507.1), Glma: *Glycine max* (NP_001276240.1), Magl: *Malpighia glabra* (ABB53473.1), Acla: *Actinidia latifolia* (ACN38266.1), Roro: *Rosa roxburghii* (AGO64472.1), Cume: *Cucumis melo* (XP_008441398.1), Soly: *Solanum lycopersicum* (XP_004240924.1), Vivi: *Vitis vinifera* (XP_002282422.1), Cisi: *Citrus sinensis* (XP_006486418.1), Thca: *Theobroma cacao* (XP_007009215.1)

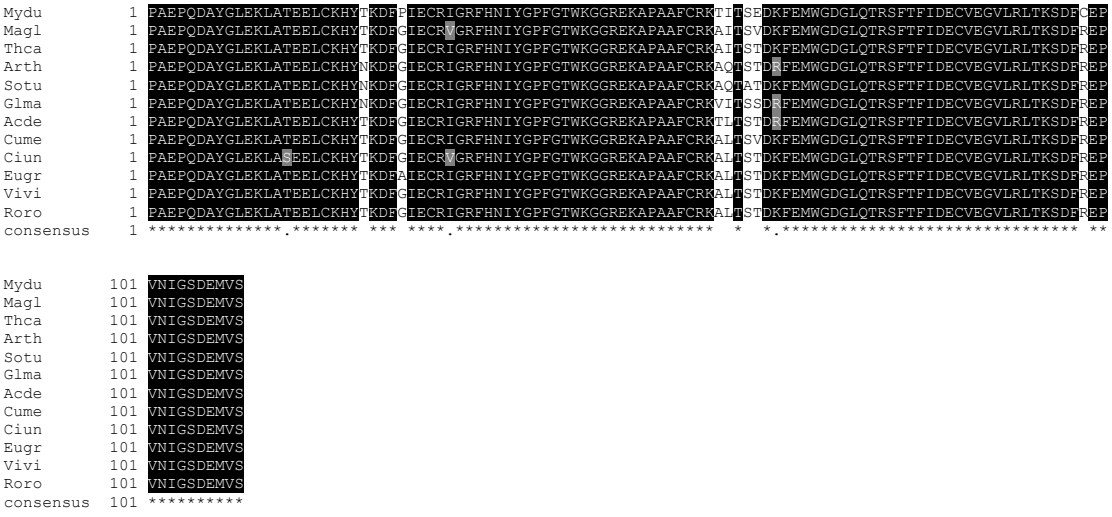

Fig. S3. Amino acid sequence alignment of GDP-D-mannose-3',5'-epimerase (GME) from *M. dubia* with sequences from other plant species. Mydu: *Myrciaria dubia* (KJ502651), Magl: *Malpighia glabra* (ABB53472.1), Thca: *Theobroma cacao* (XP_007030576.1), Arth: *Arabidopsis thaliana* (NP_198236.1), Sotu: *Solanum tuberosum* (XP_006347749.1), Glma: *Glycine max* (XP_003556467.1), Acde: *Actinidia deliciosa* (ADB85573.1), Cume: *Cucumis melo* (XP_008455931.1), Ciun: *Citrus unshiu* (ADV59924.1), Eugr: *Eucalyptus grandis* (KCW85270.1), Vivi: *Vitis vinifera* (NP_001267968.1), Roro: *Rosa roxburghii* (AGO64471.1)

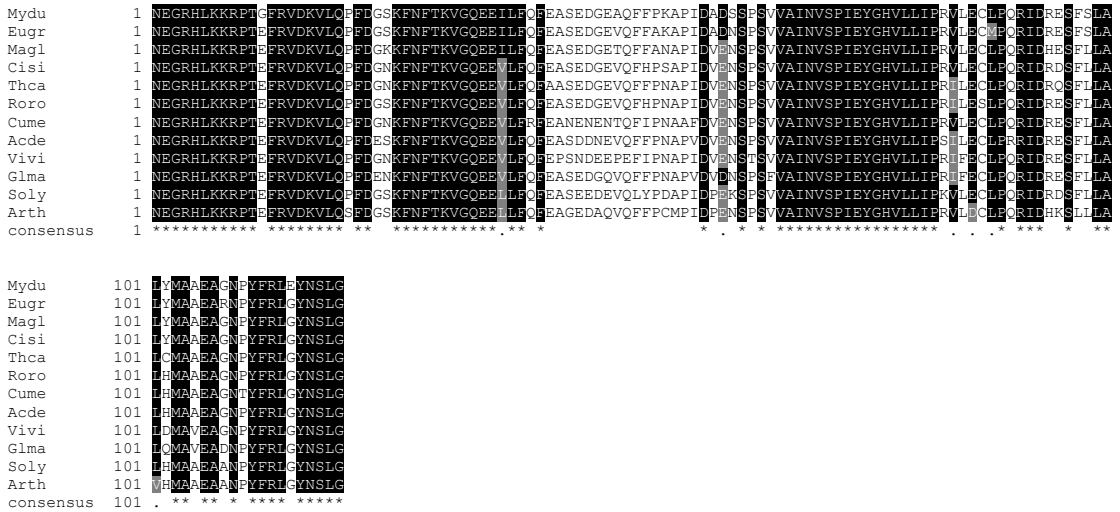

Fig. S4. Amino acid sequence alignment of GDP-L-galactose phosphorylase (GGP) from *M. dubia* with sequences from other plant species. Mydu: *Myrciaria dubia* (KJ502652), Eugr: *Eucalyptus grandis* (KCW70217.1), Magl: *Malpighia glabra* (ACG75920.1), Cisi: *Citrus sinensis* (KDO62102.1), Thca: *Theobroma cacao* (XP_007020730.1), Roro: *Rosa roxburghii* (AGO64470.1), Cume: *Cucumis melo* (XP_008447718.1), Acde: *Actinidia deliciosa* (ADB85572.1), Vivi: *Vitis vinifera* (XP_002278339.1), Glma: *Glycine max* (XP_003519588.1), Soly: *Solanum lycopersicum* (NP_001266145.1), Arth: *Arabidopsis thaliana* (NP_567759.1)

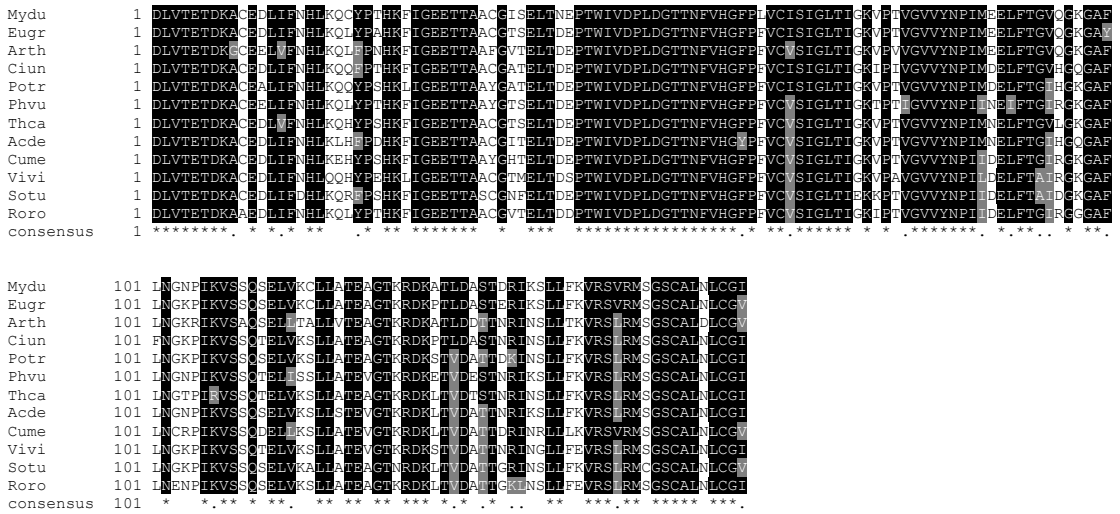

Fig. S5. Amino acid sequence alignment of L-galactose-1-phosphate phosphatase (GPP) from *M. dubia* with sequences from other plant species. Mydu: *Myrciaria dubia* (KJ502653), Eugr: *Eucalyptus grandis* (KCW75164.1), Arth: *Arabidopsis thaliana* (NP_186936.1), Ciun: *Citrus unshiu* (ADV59926.1), Potr: *Populus trichocarpa* (XP_002307895.2), Phvu: *Phaseolus vulgaris* (CAX94844.1), Thca: *Theobroma cacao* (XP_007026279.1), Acde: *Actinidia deliciosa* (AAV49506.1), Cume: *Cucumis melo* (XP_008463619.1), Vivi: *Vitis vinifera* (XP_002267760.2), Sotu: *Solanum tuberosum* (XP_006352350.1), Roro: *Rosa roxburghii* (ADK27705.1)

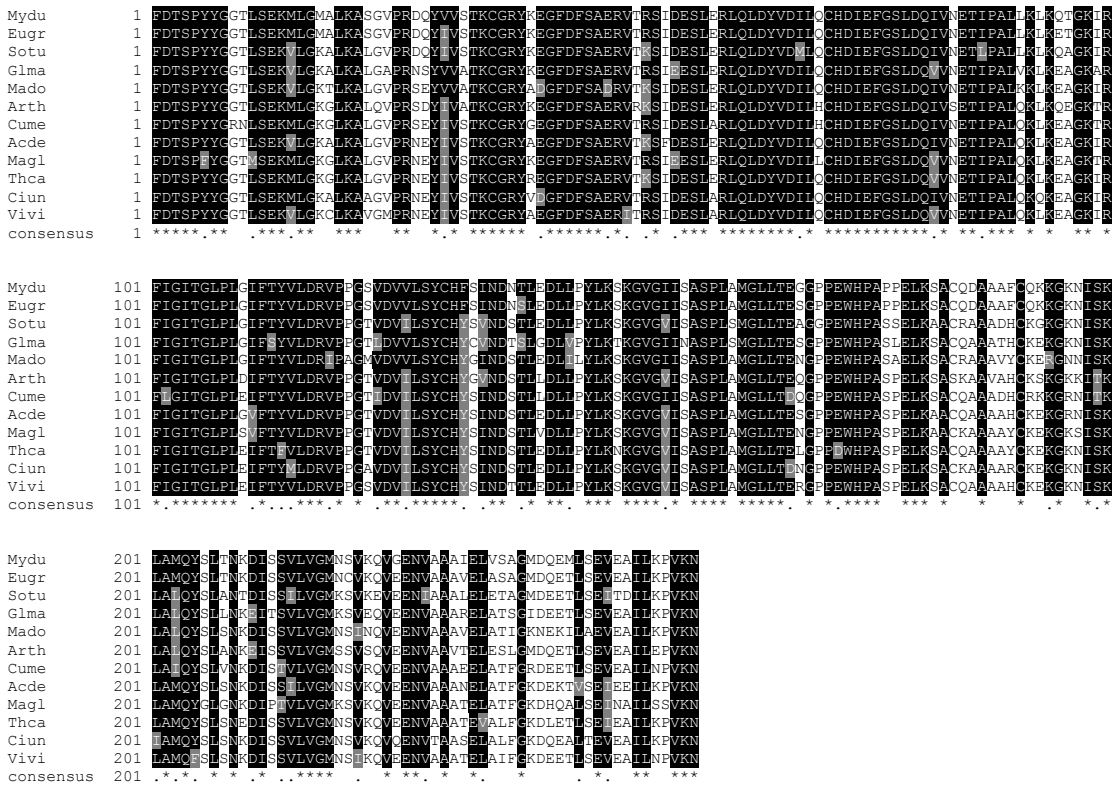

Fig. S6. Amino acid sequence alignment of L-galactose dehydrogenase (GDH) from *M. dubia* with sequences from other plant species. Mydu: *Myrciaria dubia* (KJ502654), Eugr: *Eucalyptus grandis* (KCW63489.1), Sotu: *Solanum tuberosum* (XP_006361976.1), Glma: *Glycine max* (NP_001241241.1), Mado: *Malus domestica* (NP_001281014.1), Arth: *Arabidopsis thaliana* (NP_195093.1), Cume: *Cucumis melo* (XP_008455112.1), Acde: *Actinidia deliciosa* (AAO18639.1), Magl: *Malpighia glabra* (ACG75921.1), Thca: *Theobroma cacao* (XP_007041459.1), Ciun: *Citrus unshiu* (ADV59927.1), Vivi: *Vitis vinifera* (NP_001267942.1)

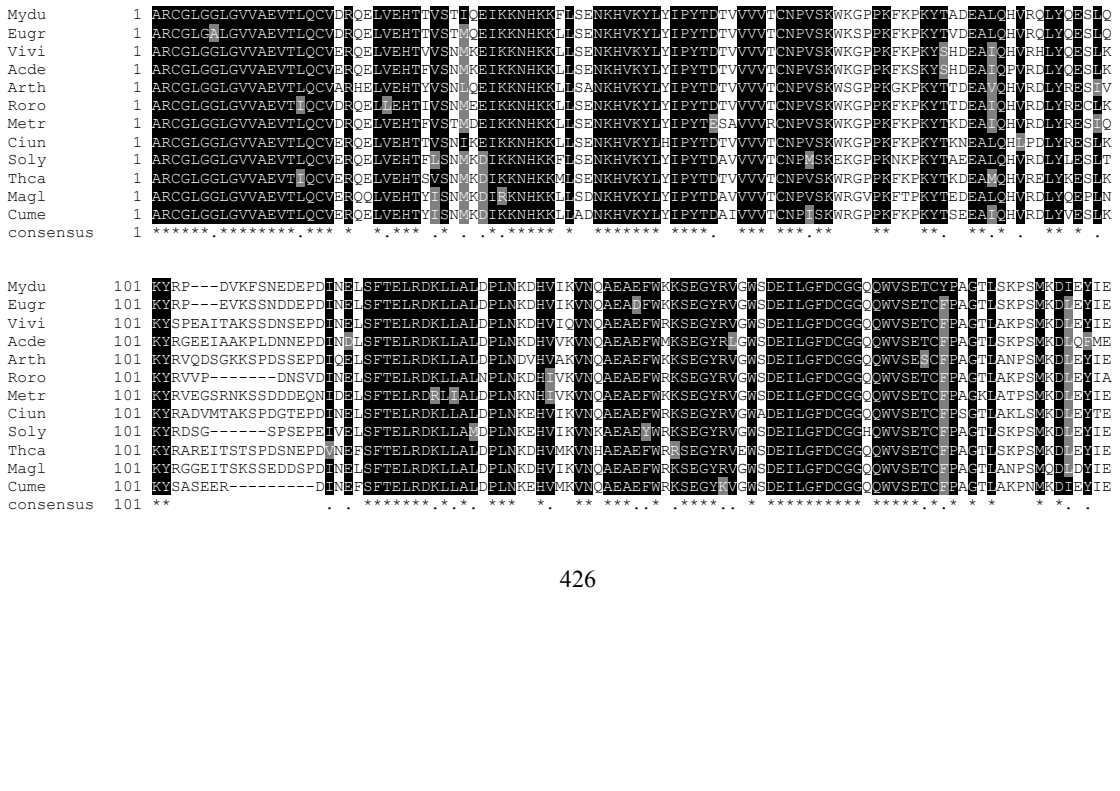

Fig. S7. Amino acid sequence alignment of L-galactono-1,4-lactone dehydrogenase (GLDH) from *M. dubia* with sequences from other plant species. Mydu: *Myrciaria dubia* (KJ502655), Eugr: *Eucalyptus grandis* (KCW90072.1), Vivi: *Vitis vinifera* (ACN59217.1), Acde: *Actinidia deliciosa* (ADB85575.1), Arth: *Arabidopsis thaliana* (BAA95212.1), Roro: *Rosa roxburghii* (AAT67210.1), Metr: *Medicago truncatula* (AES60484.2), Ciun: *Citrus unshiu* (ACO92659.1), Soly: *Solanum lycopersicum* (NP_001234603.1), Thca: *Theobroma cacao* (XP_00728561.1), Magl: *Malpighia glabra* (ACG75919.1), Cume: *Cucumis melo* (NP_001284400.1)