

**Indian Farmer**

Volume 13, Issue 09, 2026, Pp. 506-512

Available online at: www.indianfarmer.net

ISSN: 2394-1227 (Online)

Original Article

Integrative Transcriptomics and Metabolomics Reveal Candidate Genes Regulating Antioxidant and Vitamin C accumulation in Amla (*Phyllanthus emblica* L.)

Satyanshu Singh¹, S K Verma² and Khusboo Arya¹¹Ph.D. Research Scholar, Department of Fruit Science²Associate Professor, Department of Fruit Science

Acharya Narendra Deva University of Agriculture and Technology Kumarganj, Ayodhya – 224229

*Corresponding Author: abhishekagro1999@gmail.com

Received: 16/09/2026

Accepted on: 23/09/2026

ABSTRACT

Amla (*Phyllanthus emblica* L.) is an important medicinal fruit crop with remarkable antioxidant potential and extremely rich vitamin C levels. The diversity of bioactive compounds, such as emblicanin A, emblicanin B, ascorbic acid, phenolic acids, flavonoids and tannins in the fruit is major impact on its nutritional, pharmaceutical and therapeutic values. Despite the large number of biochemical investigation studies for characterization of phytochemical composition in amla, there are yet unanswered questions regarding molecular mechanisms driving accumulation of antioxidants and vitamin C accumulation. A decade of advance in high-throughput technologies for capture, sequencing and analysis of genomic DNA, and mass spectrometry-based metabolomics, have opened the door to novel integrated multi-omics strategies for dissecting complex metabolic traits. Integration of transcriptomic and metabolomic information gives a systems level understanding of the gene expression patterns, metabolic pathways and regulatory networks involved in accumulation of antioxidants. Potential molecular targets for ascorbate content improvement are key genes in vitamin C biosynthesis and recycling such as GDP-mannose pyrophosphorylase (GMP), GDP-mannose epimerase (GME), GDP-L-galactose phosphorylase (GGP), L-galactono-1,4-lactone dehydrogenase (GalLDH), monodehydroascorbate reductase (MDHAR), and dehydroascorbate reductase (DHAR). Using integration of RNA sequencing, metabolite profiling, weighted gene co-expression network analysis (WGCNA) and gene-metabolite correlation networks these candidate regulatory genes and metabolic hubs can be identified. They will be further validated as to their roles by functional validation using qRT-PCR, biochemical assays, and genome editing experiments. This multi-omics approach will be helpful for molecular breeding, metabolic engineering and development of improved cultivars of amla with antioxidant and/or nutraceutical properties.

Keywords : Transcriptomics, Metabolomics, Multi-omics integration; RNA sequencing, LC-MS/MS, Candidate genes, Molecular breeding

INTRODUCTION

Medicinal plants are important sources of bioactive compounds that find applications in the pharmaceutical industry, functional foods, nutraceuticals and also in human health promotion. Amla (*Phyllanthus emblica* L; family Phyllanthaceae) is a medicinal fruit crop which occupies a unique status among others for its extremely rich vitamin C and considerable antioxidant activity. Amla has immunomodulatory, anti-inflammatory, hepatoprotective, and antioxidant properties due to its wide range of bio-active compounds and has been used in traditional Ayurvedic medicine. The fruit contains a high amount of ascorbic acid, emblicanin A and B, gallic acid, ellagic acid, flavonoids, tannins and other phenolics. In contrast to other fruit species in which antioxidant activity has been related to a dominant metabolite, amla has a synergistic mechanism in that multiple bioactive compounds contribute to the antioxidant activity.

As a key player in the plant redox regulation pathways, ascorbic acid is involved in the detoxification of reactive oxygen species (ROS), in the regulation of photosynthesis, in stress tolerance and in the protection of cells. But, although very well characterized from a biochemical point of view, the molecular mechanisms underlying vitamin C and antioxidant accumulation in amla remain to be poorly understood. These traits are controlled by complicated interactions between structural genes, transcription factors, metabolic pathways, developmental processes and even environmental factors. Thus, combined transcriptomic and metabolomic analyses offer a valuable tool to elucidate metabolic regulation through correlating the accumulation of metabolites to the expression of genes. These multi-omics analyses will help to uncover the genetic networks, biochemical pathways, and candidate genes underlying antioxidant and vitamin C accumulation in amla.

Transcriptomic Regulation of Vitamin C Biosynthesis in Amla

Transcriptomics is a promising tool to reduce the number of genes of interest that can be linked to complex metabolic traits and nutritional quality in fruit crops. Through RNA sequencing, genome wide gene expression profiles can be obtained which allows to analyse the differentially expressed genes (DEGs) associated to fruit development, stress and antioxidant metabolism. Comparative transcriptomics of *Phyllanthus emblica* cultivars or developmental stages could contribute to understanding the difference in the antioxidant activity and accumulating the vitamin C. In this context, common tasks of a typical RNA-seq pipeline consist of RNA extraction, sequencing, quality control, genome mapping, transcript assembly, differential expression analysis, and pathway enrichment.

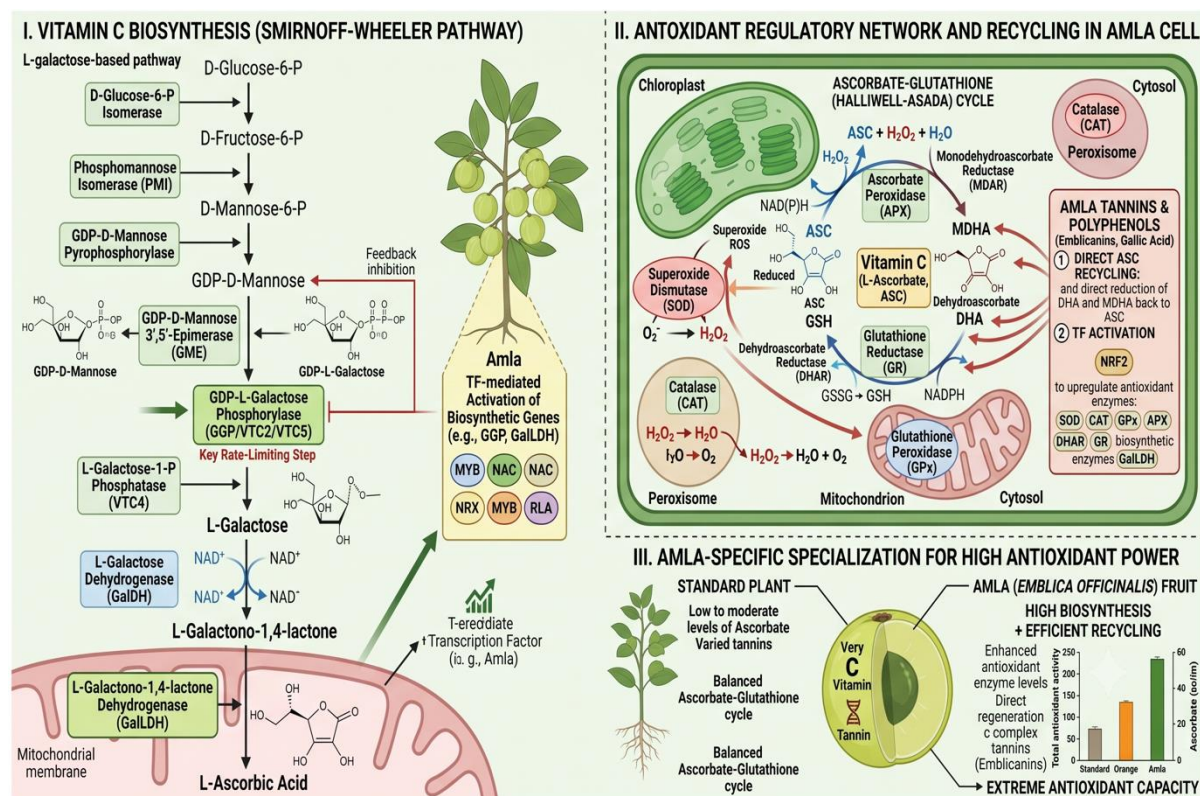


Figure : Vitamin C Biosynthesis and Antioxidant Regulatory network

Identifying genes controlling ascorbic acid biosynthesis is a major focus of transcriptomic research in amla. The L-galactose pathway is believed to be the major pathway of vitamin C biosynthesis in plants. GDP-mannose pyrophosphorylase (GMP) generates GDP-mannose and GDP-mannose epimerase (GME) generates intermediates towards GDP-L-galactose. It is known that a big regulatory and rate-limiting step is the catalysis of GDP-L-galactose phosphorylase (GPP) and changes in this expression can have a great effect on vitamin C content. L-galactono-1,4-lactone dehydrogenase (GalLDH) catalyses the final step to ascorbic acid. Besides this ascorbate recycling by monodehydroascorbate reductase (MDHAR) and dehydroascorbate reductase (DHAR), the vitamin C concentration on the cellular level is assured. Therefore, transcriptomics could be useful for identifying molecular targets to enhance the antioxidant quality of Amla.

Gene	Encoded enzyme/protein	Pathway	Biological function
GMP	GDP-mannose pyrophosphorylase	L-galactose pathway	Formation of vitamin C precursor
GME	GDP-mannose epimerase	L-galactose pathway	GDP-L-galactose production
GPP	GDP-L-galactose phosphorylase	L-galactose pathway	Rate-limiting regulation of ascorbate synthesis
GPP	L-galactose phosphatase	L-galactose pathway	Intermediate conversion

GalLDH	L-galactono-1,4-lactone dehydrogenase	Final biosynthesis step	Direct vitamin C production
MDHAR	Monodehydroascorbate reductase	Recycling pathway	Regeneration of reduced ascorbate
DHAR	Dehydroascorbate reductase	Recycling pathway	Maintains antioxidant pool
APX	Ascorbate peroxidase	ROS defence	Hydrogen peroxide detoxification
CAT	Catalase	ROS defence	Oxidative stress regulation
SOD	Superoxide dismutase	ROS defence	Superoxide radical conversion

Table 1. Candidate Genes Regulating Vitamin C Biosynthesis and Antioxidant Defence in Amla

Antioxidant Defence System and Regulatory Transcription Factors

The antioxidant defence mechanisms are closely linked to the accumulation of vitamin C. ROS are non-customized by-products of photosynthetic and respiratory processes, as well as under environmental stress experienced continuously by plants. A certain amount of ROS production is a mechanism that is used to signal, but too much will bring about oxidative damage.

An antioxidant system is comprised of non-enzymatic and enzymatic components. Superoxide radicals are scavenged by the action of superoxide dismutase (SOD) to form hydrogen peroxide, and then hydrogen peroxide is scavenged by catalase (CAT) and ascorbate peroxidase (APX). The ascorbate–glutathione cycle is one of the most important courses of antioxidant recycling which involve the activities of MDHAR, DHAR and glutathione reductase in recycling active antioxidant molecules.

There are multiple genes that are involved in the metabolism of antioxidants, whose expression can be regulated by transcription factors. The biosynthesis of phenols (phenylpropanoids and flavonoids) is regulated by MYB transcription factors. WRKY proteins are involved in signalling during conditions of stress and NAC and bZIP transcription factors are involved in fruit maturation and redox control.

Metabolomic Profiling of Antioxidant Compounds in Amla

These regulatory proteins are important molecular targets involved in understanding accumulation of antioxidants. Metabolomics offers unequivocal data on biochemical composition, and as such is fundamental for comprehending nutrition traits. Comprehensive identification of metabolites associated with antioxidant activity is possible using advanced platforms like LC-MS/MS, GC-MS and NMR.

The phenolics, flavonoids, tannins and vitamin C derivatives are especially suitable for detection using LC-MS/MS. GC-MS can provide information on primary metabolites while NMR can

provide reproducible metabolic fingerprinting. Amla contains the following major antioxidants metabolites: ascorbic acid, emblicanin's, gallic acid, ellagic acid and derivatives of flavonoids.

Metabolite	Chemical category	Biological significance
Ascorbic acid	Vitamin antioxidant	Major antioxidant and nutritional compound
Emblicanin A/B	Hydrolysable tannins	Enhances antioxidant stability
Gallic acid	Phenolic acid	Free radical scavenging
Ellagic acid	Polyphenol	Oxidative protection
Quercetin derivatives	Flavonoids	ROS regulation
Kaempferol derivatives	Flavonoids	Antioxidant defence

Table 2. Major Antioxidant Metabolites of Amla and Their Biological Importance

The key point of a multi-omics approach is that it allows to correlate what happens at the gene level to what happens at the metabolic level. Integration of transcriptomic data gives stronger evidence to determine whether a gene contributes to metabolite accumulation as transcriptomic data alone does not provide definitive evidence of this.

Gene-metabolite correlation analysis involves the identification of genes whose accumulation trends closely mirror the accumulation trends seen for metabolites. Key computation techniques are:

- Differential expression analysis
- KEGG pathway enrichment
- Weighted gene co-expression network analysis (WGCNA)
- Pearson correlation networks
- Principal component analysis

WGCNA makes use of clusters of genes with comparable expression, and connects them with antioxidant traits. The genes at the central positions are classified as hub genes. For instance, the concerted change of GGP expression with a high vitamin C content suggests a possible regulatory function of GGP in the accumulation ascorbate. Likewise, MYB genes that showed significant association with flavonoids metabolites could be regulatory factors for the production of flavonoids.

Functional Validation of Candidate Genes

Multi-omics approaches offer strong prediction of functional validation is needed to determine biological significance. Analysing gene expression at RNA level by quantitative real-time PCR (qRT-PCR) confirms the findings in many cases that were obtained from RNA-seq expression patterns. Candidate genes like, GMP, GGP, GalLDH, MDHAR, DHAR, APX, CAT and SOD, may be analysed in various tissues and developmental stages etc.

Activity assays of enzymes: provide evidence of gene expression related to biochemical function. More sophisticated techniques include the over-expression of genes, RNA interference and

CRISPR/Cas genome editing. Gene over-expression of genes like GGP, or GalLDH could improve the accumulation of vitamin C and silencing experiments could show if those genes are necessary for antioxidant production.

Approach	Information generated	Application
RNA sequencing	Differentially expressed genes	Candidate gene identification
LC-MS/MS metabolomics	Differential metabolites	Antioxidant compound profiling
WGCNA	Co-expression modules	Hub gene identification
Gene-metabolite correlation	Regulatory networks	Candidate gene prediction
qRT-PCR and enzyme assays	Functional confirmation	Biological validation

Table 3. Multi-Omics and Functional Validation Approaches

Applications in Molecular Breeding and Future Perspectives

The discovery of antioxidant regulatory genes can open the possibility for molecular improvement of amla. Considering that antioxidant traits are quantitative and subjected to environmental influences, traditional approach would require long durations. Using molecular markers, multi-omics-assisted breeding can increase the speed, i.e. reduce the amount of time spent, necessary for the selection of better genotype.

Phenols and vitamins could be targets for the following genes:

- Marker-assisted selection
- Genomic selection
- Metabolic engineering
- Functional food development

The combination of multi-omics and artificial intelligence, single-cell transcriptomics, and spatial metabolomics should be considered for future research. The possibility to predict important regulatory genes using AI-based approaches could alleviate the burden of performing large-scale omics analyses, while spatial metabolomics could provide information on where antioxidants are present in different tissues within fruits.

CONCLUSION

Amla is a rare medicinal fruit crop with supreme anti-oxidants and Vitamin C accumulation ability. But, to control these traits is hard to do, and needs high-end systems biology tools. Combining data from transcriptomics and metabolomics gives a comprehensive approach to identify putative antioxidant metabolome regulating genes. Changes in gene expression analysed by transcriptomic and biochemical changes analysed by metabolomic include those linked to ascorbate biosynthesis, ROS regulation and secondary metabolism.

The genes GMP, GME, GGP, GalLDH, MDHAR, DHAR and transcription factors like MYB and WRKY are important potential target genes for research on antioxidant accumulation in amla. Again, the synergy of multi-omics technologies, computational biology and functional genomics will speed up the process of development of innovative high-value amla cultivars and therefore ensure sustainable production of antioxidant-rich functional foods and therapeutic resources.