A GENOME-WIDE ANALYSIS OF THE *GALACTINOL SYNTHASE* GENE FAMILY IN BANANA (*Musa acuminata*)

Fernanda Dolcimasculo[1], Alessandra Ferreira Ribas[2], Luiz Gonzaga Esteves Vieira[2], Tiago Benedito dos Santos[1,2]

[1]Universidade do Oeste Paulista – UNOESTE, Especialização Lato Senso em Biotecnologia, Presidente Prudente, SP. [2]Universidade do Oeste Paulista – UNOESTE, Programa de Pós-Graduação em Agronomia, Presidente Prudente, SP. E-mail: dossantos@unoeste.com

ABSTRACT

Galactinol synthase (GolS) is the enzyme that catalyzes the first step of the biosynthesis of the raffinose family oligosaccharides (RFOs), and is involved in many biological processes in plants. In the present study, four putative *GolS* genes were identified in the *Musa acuminata* genome. We further characterized these *MaGolS* genes in terms of protein length, molecular weight, theoretical isoelectric point and 3D protein structure. Genomic organization revealed that most *MaGolS* genes have four exons. The conserved motifs were identified, demonstrating high group-specificity of all MaGolS proteins. Multiple sequence alignment showed that the APSAA typical domain is present in all GolS proteins. Comparative phylogenetic analysis of the MaGolS proteins revealed three distinct groups. These data provide insight to support new studies adressing the role of *GolS* genes in this important fruit species.

Key words: banana; i*n silico*; raffinose family oligosaccharides; *GolS* gene.

ANÁLISE GENÔMICA DA FAMÍLIA DE GENES DA *GALACTINOL SINTASE* EM BANANA (*Musa acuminata*)

RESUMO

A galactinol sintase (GolS) é a enzima que catalisa o primeiro passo da biossíntese dos oligossacarídeos da família da rafinose (OFRs) e está envolvida em muitos processos biológicos nas plantas. No presente estudo, quatro genes *GolS* foram identificados no genoma de *Musa acuminata*. Esses genes *MaGolS* foram caracterizados em termos de comprimento de proteína, peso molecular, ponto isoelétrico teórico e estrutura 3D de proteína. A organização genômica revelou que a maioria dos genes *MaGolS* possui quatro éxons. Os motivos conservados foram identificados, demonstrando alta especificidade de grupo de todas as proteínas MaGolS. O alinhamento múltiplo das sequências mostrou que o domínio típico de APSAA está presente em todas as proteínas GolS. A análise filogenética comparativa das proteínas MaGolS revelou três grupos distintos. Estes dados fornecem *insights* para apoiar novos estudos sobre o papel dos genes *GolS* nesta importante espécie frutífera.

Palavras-chave: banana; *in silico*; oligossacarídeos da família da rafinose; gene *GolS*.

INTRODUCTION

Banana (family Musaceae, order Zingiberales) is considered one of the most preferred fruits by the world population, besides being an important crop in developing countries (LESCOT, 2014). The majority of the commercial bananas is derived from the interspecific cross between *Musa acuminata* Colla (2n = 2× = 22, A genome) and *M. balbisiana* Colla (2n = 2× = 22, B genome) (HESLOP-HARRISON; SCHWARZACHER, 2007; SIMMONDS; SHEPERD, 1955; PRICE, 1995; SALOMÃO; SIQUEIRA, 2015). It is believed that the *M. acuminata* species is divided into six to nine subspecies (banksii, burmannica, malaccensis, microcarpa, zebrina, burmannicoïdes, truncata, siamea, and errans), which possibly diverged following geographical isolation in Southeast Asian continental islands (DANIELLS et al., 2001; PERRIER et al., 2009).

Among the plant carbohydrates, the family of raffinose oligosaccharides (RFOs) has been poorly studied in monocotyledonous crop plants, although it is well understood in different dicotyledonous species as it acts in response to different biological stimuli and signals. RFOs are also essential for plant growth and are described to play an important role in abiotic stress tolerance (DOS SANTOS et al., 2011; 2015). Galactinol synthase (GolS, EC 2.4.1.123), the key enzyme in the biosynthesis of the RFOs, integrate a subfamily of the glycosyltransferase 8 (GTs; EC 2.4.x.y.), which is composed of enzymes that participate in the biosynthesis of sugars involved in signaling pathways, cell structure composition and energy reserve (SENGUPTA et al., 2012; 2015). The precursor molecule of the biochemical pathway of the RFOs, galactinol, results from the catalytic activity of the GolS, which uses UDP-D-galactose and a *myo*-inositol molecule to generate its product in the cytosol (SCHNEIDER; KELLER, 2009). This enzyme has a key regulatory process in the metabolism of sugars in plants, and in the partitioning of the carbon between sucrose-dependent synthesis pathways and RFOs (NISHIZAWA et al., 2008).

In recent years, with the growing availability of information generated by genome sequencing in plants, many studies have been conducted for identifying and understanding the functional significance of inumerous genes. *GolS* genes have been the subject of several studies in many plant species, and are genes that are up-regulated in reponse to multiple developmental and environmental responses (DOS SANTOS et al., 2015; KANNAN et al., 2016; LIU et al., 2016; JANG et al., 2018; CHU et al., 2018). In an effort to accelerate the genomic studies in banana, a double haploid (DH-Pahang) had its genome sequenced, and a platform was made available for the characterization of its genes (D'HONT et al., 2012; DROC et al., 2013). Until now, there have been no studies on the identification and characterization of the *GolS* gene family in banana. Many allelic variants of the *GolS* gene have been characterized and described in the literature which are mainly involved in abiotic stress. For example, in *Arabidopsis thaliana* seven *GolS* isoforms (classified as *AtGolS1 to 7*) were identified. Among them, *AtGolS1, 2* and *3* are reported to be involved in abiotic stress responses (TAJI et al., 2002). dos Santos et al. (2011; 2015) also observed that three isoforms in *Coffea arabica* L. and *C. canephora* are involved in the mechanisms of abiotic stress tolerance.

In this study, based on the recent availability of genome of the *M. acuminata* (DROC et al., 2013), we employed bioinformatic analysis to characterize the *GolS* genes. Here, we analyzed their phylogenetic relationship with homologous genes in other plant species, gene structure, protein motifs and 3D structure. The identification and comprehensive investigation of the *MaGolS* genes in banana could be a useful tool for future biotechnological research on genetic improvement of banana.

MATERIAL AND METHODS

Identification and classification of *MaGolS*

The *GolS* genes sequences (nucleotides, peptides and CDS) of *M. acuminata* were dowloaded from the Banana Genome Hub (DROC et al., 2013). A total of four gene sequences were retrieved from plattaform and are listed in Table 1. Each putative *MaGolS* gene was individually confronted with sequences deposited in the NCBI database (BlastX and BlastP algorithm; ALTSHUL et al., 1990) in order to verify the specificity of the annotated sequences. The *MaGolS* genes were used as queries in the TAIR10 plattaform (http://www.arabidopsis.org; BlastP tool) to finding the best hit to the *Arabidopsis* reference database (Table 1). For nomenclature of the identified genes, the *MaGolS* genes were numbered according to their positions in the chromosomes 1–11.

Analysis of MaGolS proteins

Detailed information about the MaGolS proteins identified in the present study, including the physical and chemical properties, number of amino acids (aa), molecular weight (MW) and theoretical pI (isoelectric point) were calculated using the ExPASy ProtParam tool (http://web.expasy.org/protparam/).

Gene structure and motifs of *MaGolS*

Characterization of gene structure (exon/intron) of *MaGolS* genes were obtained by mapping the cDNA to DNA sequences using the Gene Structure Display Server2.0 (GSDS; http://gsds.cbi.pku.edu.cn/; HU et al., 2015).

The subcellular localizations were predicted in the Plant-mPLoc server (http://www.csbio.sjtu.edu.cn/bioinf/euk-multi-2/ and http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/; CHOU, SHEN, 2007; CHOU, SHEN, 2010). MaGolS protein conserved motifs were analyzed with Multiple Expectation Maximization for Motif Elicitation (MEME; BAILEY et al., 2006) software, with the following parameter: minimal and maximal motif width were set to 6 and 100 amino acids and the number of motifs to find was set to 15.

Multiple sequence alignment and phylogenetic analysis

Mutiple sequence alignment was performed on the protein sequences of *M. acuminata* using CLC Main Workbench v.8 program with default parameters. For investigate the phylogenetic relationship of the *MaGolS* genes in banana, representative protein sequences of different plant species were retrived from NCBI: *A. thaliana* (NP_1822401.1 – AtGolS1; NP_176053.1 – AtGolS2; NP_172406.1 – AtGolS3; NP_176250.1 – AtGolS4; NP_197768.1 – AtGolS5; NP_567741.2 – AtGolS6; NP_176248.1 – AtGolS7), *Zea mays* L. (AAQ07248.1 – ZmGolS1; AAQ07249.1 – ZmGolS2; AAQ07250.1 – ZmGolS3). *Brachypodium distachyon* (BdGolS1 – Bradi1g64120; BdGolS2 – Bradi1g17200) sequences were downloaded from the Phytozome v.12 (http://phytozome.jgi.doe.gov/pz/portal.html; GOODSTEIN et al., 2012) database. The MEGA7.0 program was utilized to construct the phylogenetic tree using the Neighbor-Joining method, p-distance substitution model, pairwise deletion, with bootstrap value of 1.000 independent replicates (KUMAR et al., 2016).

3D Structure

The three-dimensional (3D) structure prediction of *MaGolS* genes were generated by Phyre2 protein-modeling server (www.sbg.bio.ic.ac.uk/*phyre2). Phyre2 is a tool used in the analysis and prediction of protein structures, in addition to function and protein mutations.

RESULT AND DISCUSSION

GolS genes have already been surveyed in other species, such as *A. thaliana* (TAJI et al., 2002), *Salvia miltiorrhiza* (WANG et al., 2012), *C. arabica* and *C. canephora* (DOS SANTOS et al., 2011; 2015), *Camellia sinensis* (ZHOU et al., 2017), *Malus* × *domestica* (FALAVIGNA et al., 2018), *Sesamum indicum* L. (YOU et al., 2018) and *Glycine max* (CHU et al., 2018). In this study, we performed an *in silico* analysis of the four *GolS* genes identified in the *M. acuminata* genome.

The screening of the *MaGolS* sequences using BlastP searches, excluding all non-redundant sequences that did not present the GT8 domain (PF01501), indicated only four putative galactinol synthase genes (Table 1). The number predicted for *GolS* genes in the genomes of different plant species is variable. As mentioned earlier, seven *GolS* genes were reported in *A. thaliana* (TAJI et al., 2002). In a study on the evolutionary diversification of GolS family in *M. domestica*, eight genes were identified (referred to as *MdGolS1* to *8*; FALAVIGNA et al., 2018). You et al. (2018), studying the expression of genes involved in raffinose accumulation, identified seven *GolS* genes in *S. indicum* L., some of which are involved in abiotic stress responses.

The allelic variants found here were composed of 328 (*MaGolS1*), 327 (*MaGolS2*), 334 (*MaGolS3*) and 342 (*MaGolS4*) (aa) amino acid residues, respectively. Those isoforms were further characterized as to have molecular weights and pIs ranging from 36737.46 to 38099.87 and 5.08 to 8.04, respectively (Table 1). Compared to the *GolS* orthologs genes of *Arabidopsis*, the AT1G60470.1 was the best hit with 76% identity at the amino acid level to the *MaGolS1* query gene, and AT1G60480.1 with *MaGolS2*, *MaGolS3* and *MaGolS4* (Table 1).

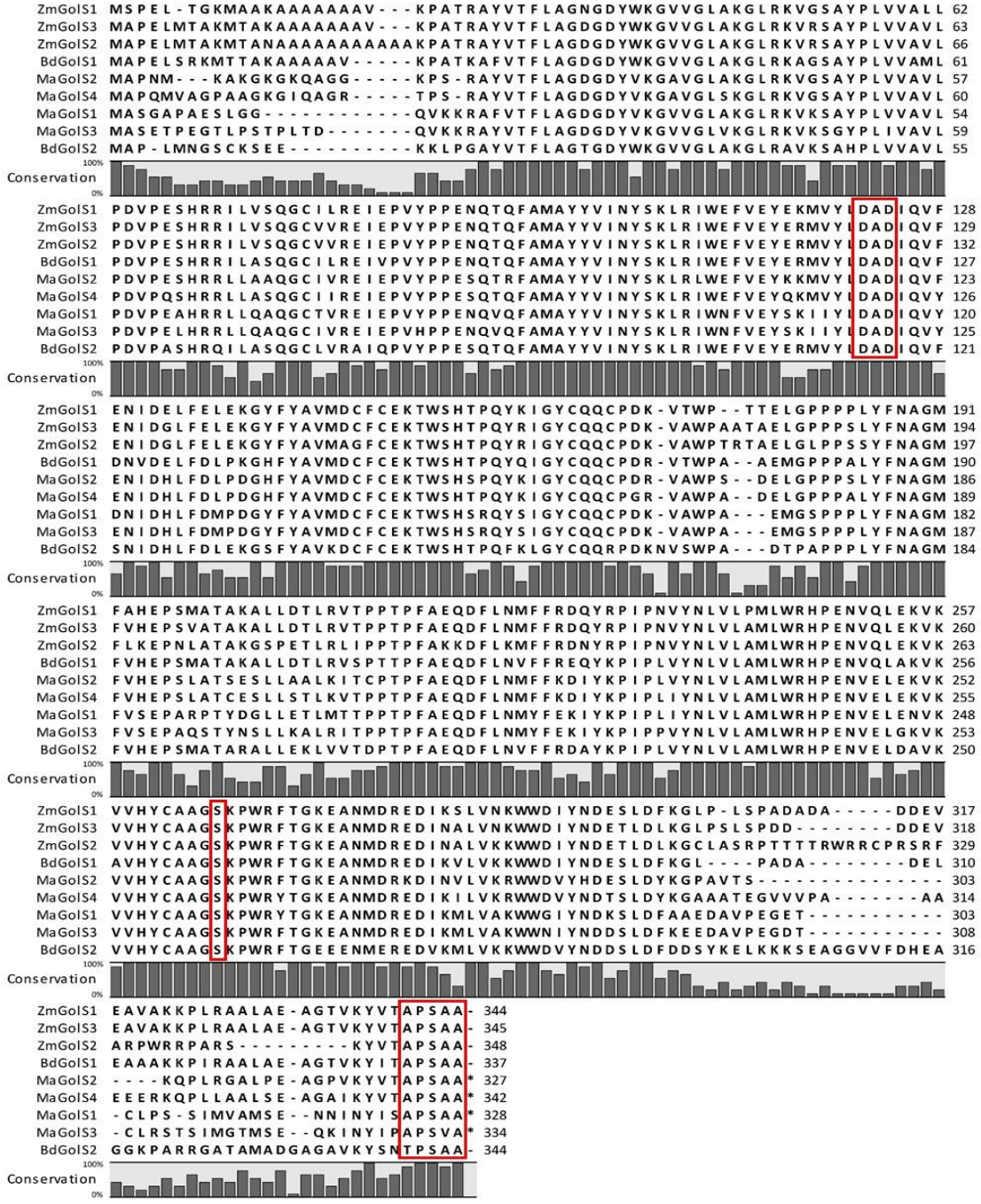
Table 1. List of *MaGolS* genes identified in *M. acuminata* with their corresponding nucleotide and protein lengths, physical-chemical characteristics of proteins, isoeletric point (pI), subcellular localization and orthologs from *A. thaliana*.

Gene name	Annotation ID	Chromossome	Lenght (bp)	ORF (aa)	Molecular weight	Theoretical pI	Subcellular localization	Ortholog in *Arabidopsis*
MaGolS1	Ma06_p29050.1	chr06:30548147..30552178	2.465	328	36977.46	5.08	Cytoplasm	AT1G60470.1
MaGolS2	Ma08_p10610.1	chr08:7765238..7766798	1.561	327	36737.46	8.04	Cytoplasm	AT2G47180.1
MaGolS3	Ma09_p06720.1	chr09:4297532..4300004	2.473	334	38015.75	5.48	Cytoplasm	AT2G47180.1
MaGolS4	Ma11_p14680.1	chr11:20393459..20394845	1.387	342	38099.87	5.85	Cytoplasm	AT2G47180.1

The predicted subcellular localization of all MaGolS proteins were found in the cytoplasm (Table 1). According to Keller (1992), GolS is an extravacuolar enzyme, probably cytosolic. Our data corroborate previous studies that have shown that GolS enzymes are localized exclusively in the cytoplasm as no plastid transit peptides were found in known GolS using different sequence-based predictors (EMANUELSSON et al., 2007).

A multiple sequence alignment analysis of the MaGolS proteins with the four putative genes is shown in Figure 1.

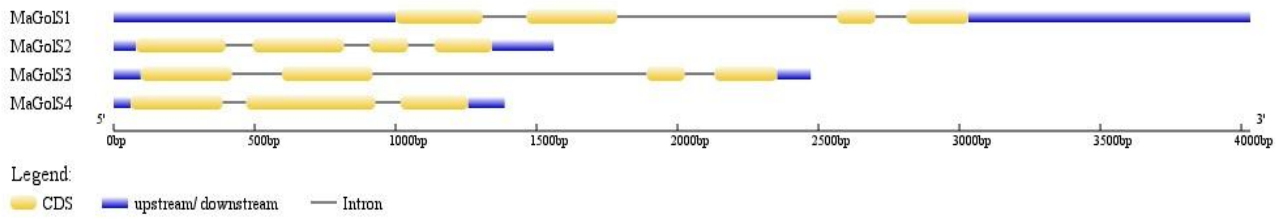
Figure 1. Deduced amino acid sequence alignment of GolS from *M. acuminata* and other monocots. The amino acid sequences used in the multialignment analysis were from *B. distachyon* (BdGolS1 – Bradi1g64120; BdGolS2 – Bradi1g17200), Z. *mays* L. (ZmGolS1 –AAQ07248.1; ZmGolS2 – AAQ07249.1; ZmGolS3 – AAQ07250.1). The conserved catalytic residues of a manganese-ligation motif (DxD), the typical serine phosphorylation site (S) and the hydrophobic pentapeptide (APSAA) located at the C-terminal are contained in the red box.



All amino acid sequence of the MaGolS protein showed the C-terminal hydrophobic pentapeptide APSAA (SPRENGER; KELLER, 2000), a conserved characteristic of this gene family, with the exception of the *MaGolS3,* in which a substitution of alanine (A) for valine (V; APSVA) in the penultimate amino acid position was observed (Figure 1). Despite the ubiquitous presence of this motif in *GolS* genes, it was not present in one isoform (*BdGolS2*) in *B. distachyon* (FILIZ et al., 2015). Based on that observation, these authors suggested that this conserved sequence may not be essential for the enzyme activity. All together, our data corroborate other studies regarding the characteristic C-terminal sequence of the *GolS* genes, the presence of conserved manganese-ligation motif (DxD) and the serine phosphorylation site (S), as observed in *C. arabica, C. canephora* and *C. sinensis* (DOS SANTOS et al., 2011, 2015; ZHOU et al., 2017).

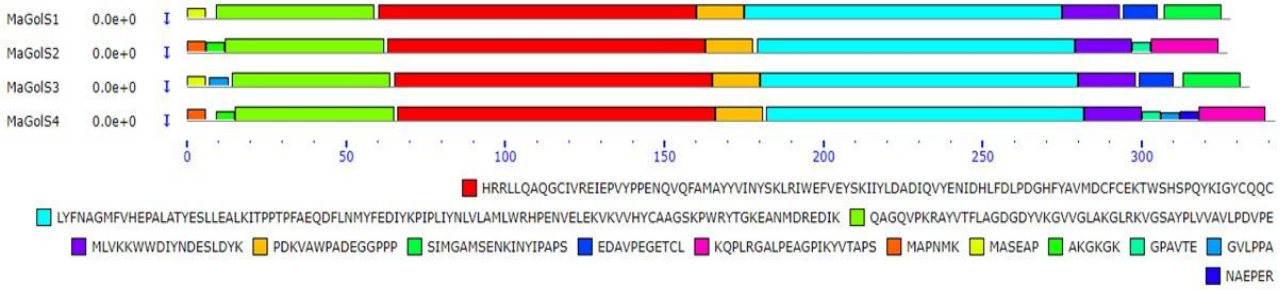
There was little variation in the number of introns (varying from 2 to 3) between the four *GolS* genes of *M. acuminata* (Figure 2).

Figure 2. Gene structure of *MaGolS* genes. The exons, introns and untranslated region (UTR) are indicated by yellow rectangles, gray lines and blue rectangles, respectivelly.



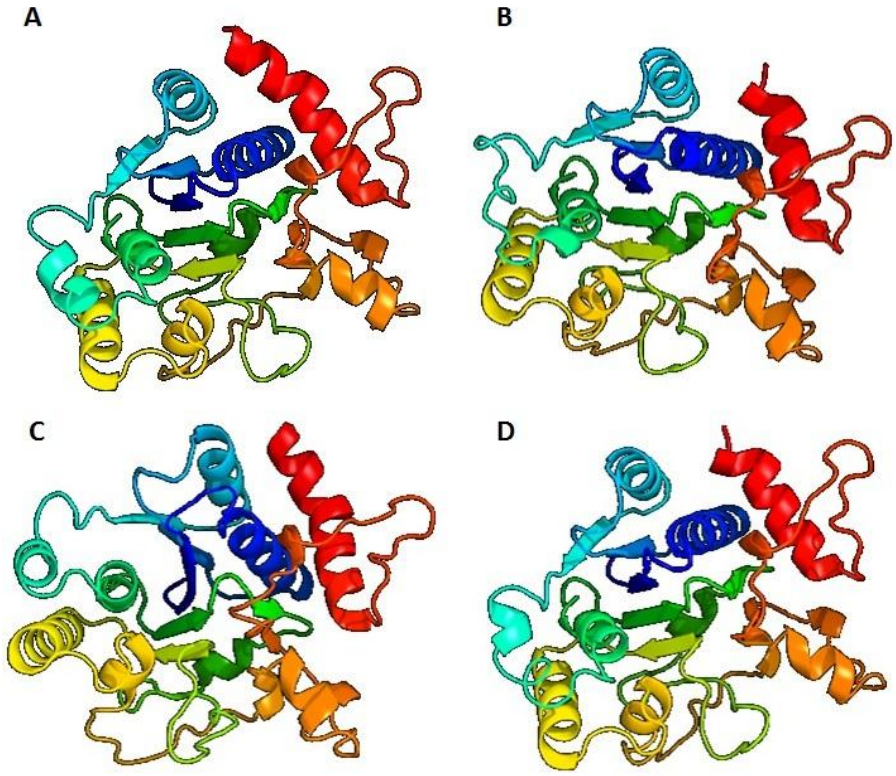
Similar gene structure were reported in other plant taxa, such as in *S. miltiorrhiza* (WANG et al., 2012), *Pisum sativum* (LAHUTA et al., 2014), *Solanum lycopersicum, B. distachyon* (FILIZ et al., 2015), *M. domestica* (FALAVIGNA et al., 2017), *Brassica napus*, *Nicotiana tabacum* (FAN et al., 2017) and *S. indicum* L. (YOU et al., 2018). As revised by Filiz et al. (2015), the variations in gene structure and number may be caused by mutations and rearrangements in *GolS* genes occurred during monocot and dicot species split.

Figure 3. Schematic representation of the conserved motifs in the MaGolS proteins. Amino acid motifs are represented by colored boxes, respectively.



The 3D structure predictions for the *MaGolS* genes were generated using the Phyre2 server in order to understand the structural properties of these genes (Figure 4).

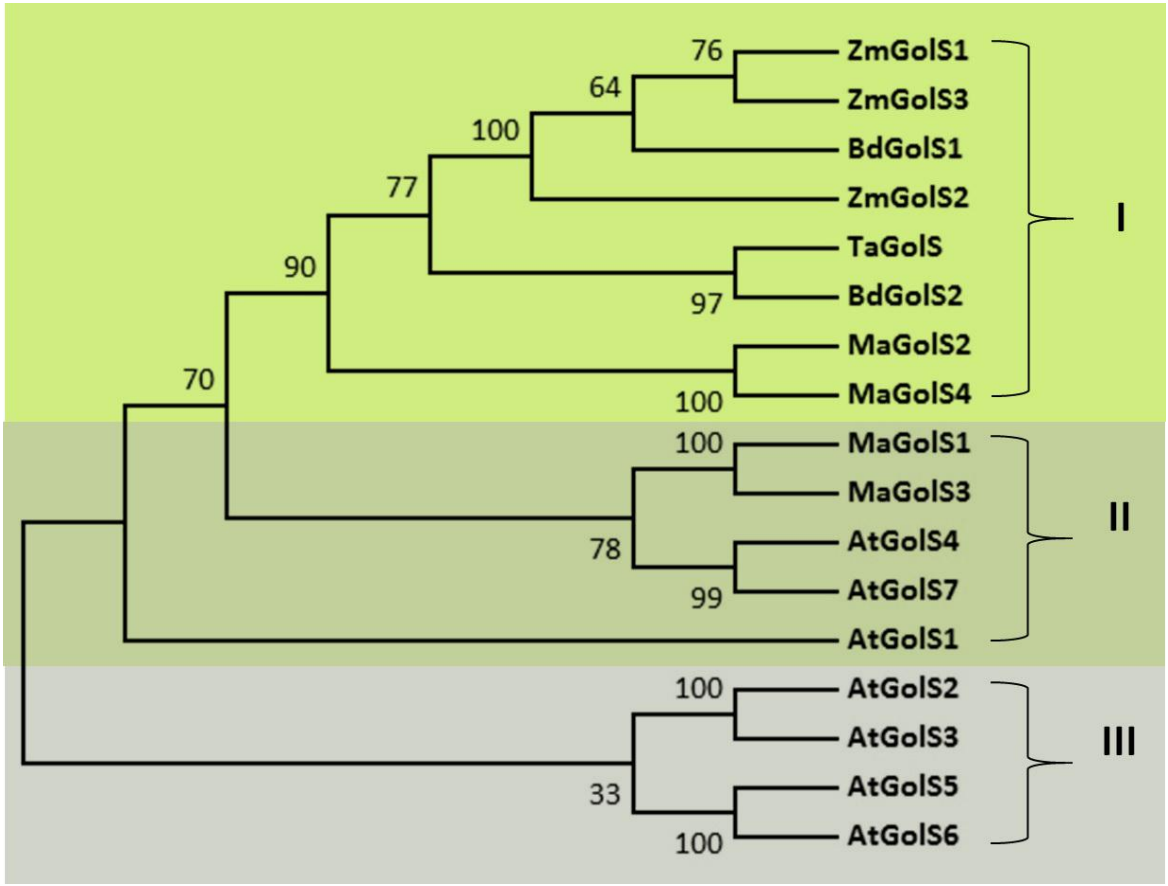
Figure 4. Predicted 3D structures of the four MaGolS proteins. **(A)** MaGolS1, **(B)** MaGolS2, **(C)** MaGolS3 and **(D)** MaGolS4.



The 3D protein models have been constructed with 100% confidence and residue coverage between the *MaGolS* genes varied from 71 to 75%. Thirteen α-helices and eigth β-strands were predicted for MaGolS1, MaGolS3 and MaGolS4; whereas eleven α-helices and ten β-strands were predicted for MaGolS2. The best template (d1ll2a_) used for the 3D structure predication belongs to the family glycogenin of the superfamily (Nucleotide-diphospho-sugar transferases) (GIBBONS et al., 2002). Our data confirm that *GolS* genes in monocot present a noticiable structural diversity; for example, ZmGolS3 of *Z. mays* contains 20 α-helices, 7 β-strands (SENGUPTA et al., 2012), *B. distachyon* BdGolS1 has 14 α-helices and 20 β-strands and BdGolS2 has 10 α-helices and 8 β-strands (FILIZ et al., 2015).

A phylogenetic tree was constructed from the complete GolS proteins using selected monocots (*Z. mays*, *B. distachyon, Triticum aestivum*) and one dicot species (*A. thaliana*) serving as an outgroup and revealed three distinct clades: GolSI, GolSII and GolSIII (Figure 5).

Figure 5. Phylogenetic tree of GolS proteins from *M. acuminata* (MaGolS) and four other species (*A. thaliana*, *Z. mays*, *B. distachyon and T. aestivum*). The sequences were aligned using ClustalW at MEGA 7.0 software and the phylogenetic tree was constructed by Neighbor-Joining method. Groups are distinguished by different color boxes.



Two MaGolS (2 and 4) were clustered into group GolSI, which included ZmGolS1, ZmGolS3, BdGolS1, ZmGolS2, TaGolS and BdGolS2. The allelic variants MaGolS1 and 3 were clustered into group GolSII, which included AtGolS4 and 7 (Figure 5). The group GolSIII contained only members from *Arabidopsis*. According to Blanc and Wolfe (2004) and reviewed by Filiz et al. (2015), 62% of duplicated genes in *Arabidopsis* presented a differential expression pattern which may lead to divergences in their function. The results demonstrated that MaGolS were arranged in pairs (MaGolS2 and MaGolS4, MaGolS1 and MaGolS3), suggesting that *MaGolS* genes may have undergone duplication along the evolution and adaptation of this species. This hypothetical paralogs of *MaGolS* genes may also hint at the possible divergence of *GolS* functions in banana.

CONCLUSION

To our knowledge, this is the first genome-wide study of the *M. acuminata GolS* gene. We identified four *MaGolS* genes that were characterized according to the amino acid residues, the conserved motifs and 3D structure of proteins and phylogenetical analysis. Results obtained from this study provide a basis for further studies aiming at the functional characterization of *MaGolS* genes in banana.

REFERENCES

ALTSCHUL, S.F.; GISH, W.; MILLER, W.; MYERS, E.W.; LIPMAN, D.J. Basic local alignment search tool. **Journal of Molecular Biology**, v.215, n.3, p.403-410, 1990. https://doi.org/10.1016/S0022-2836(05)80360-2

BAILEY, T.L.; WILLIAMS, N.; MISLEH, C.; LI, W.W. MEME: Discovering and analyzing DNA and protein sequence motifs. **Nucleic Acids Research**, v.34, n.2, p.369-373, 2006. https://doi.org/10.1093/nar/gkl198

BLANC, G.; WOLFE, K.H. Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. **Plant**

Cell, v.16, p.1667-1678, 2004. https://doi.org/10.1105/tpc.021345

CHU, H. C.; BUFFEL, M.; TIEN, D. L. Functional characterisation of a soybean galactinol synthase gene under various stress conditions. **Vietnam Journal of Science, Technology and Engineering**, v.60, n.3, p.33-36, 2018. https://doi.org/10.31276/VJSTE.60(3).33

CHOU, H.C.; SHEN, H.G. Euk-mPLoc: A fusion classifier for large-scale eukaryotic protein subcellular location prediction by incorporating multiple sites. **Journal of Proteome Research**, v. 6, p.1728 -1734, 2007. https://doi.org/10.1021/pr060635i

CHOU, H.C.; SHEN, H.G. Plant-mPLoc: a top-down strategy to augment the power for predicting plant protein subcellular localization. **PLoS ONE**, v.5, 2010. https://doi.org/10.1371/journal.pone.0011335

DANIELLS, J.; JENNY, C.; KARAMURA, D.; TOMEKPE, K. Musalogue : a catalogue of Musa germplasm: diversity in the genus Musa. **Biover Int**. p.213, 2001.

D'HONT, A.; DENOEUD, F.; AURY, JM.; BAUREN, F.C.; CARREEL, F.; GARSMEUR, O.; NOEL, B.; BOCS, S; DROC, G.; ROUARD, M.; DA SILCAA, C.; JABBARI, K.; CARDI, C.; POULAIN, J.; SOUQUET, M.; LABADIE, K.; JOURDA, C.; LENGELLÉ, J.; RODIER-GOUD, M.; ALBERT, A.; BERNARD, M.; CORREA, M.; AYYAMPALAYAM, S.; MCKAIN, M.R.; LEEBENS-MACK, J.; BURGUESS, D.; FREELING, M.; MBÉGUIÉ-A-MBÉGUIÉ, D.; CHABANNES, M.; WICKER, T.; ET AL. The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. **Nature**, v.488, p.213-217, 2012. https://doi.10.1038/nature11241.

DROC, G.; LARIVIERE. D.; GUIGNON, V.; et al. The Banana Genome Hub. **Database**, v.2013, bat035, 2013. https://doi.org/10.1093/database/bat035

EMANUELSSON, O.; BRUNAK, S.; VON HEIJNE, G.; NIELSEN, H. Locating proteins in the cell using TargetP, SignalP and related tools. **Nature Protocols**, v.2, p.953-971, 2007. https://doi.org/10.1038/nprot.2007.131

FALAVIGNA V.S.; PORTO D.D.; MIOTTO Y.E.; SANTOS H.P.D.; OLIVEIRA P.R.D., MARGIS-PINEHIRO M. Evolutionary diversification of galactinol synthases in Rosaceae: adaptive roles of galactinol and raffinose during apple bud dormancy. **Journal of Experimental Botany**, v.69, n.5 p.1247-1259, 2018. https://doi.org/10.1093/jxb/erx451

FAN, Y.; YU, M.; LIU, M.; ZHANG, R.; SUN, W.; QIAN, M.; ZHANG, K. Genome-Wide Identification, Evolutionary and Expression Analyses of the GALACTINOL SYNTHASE Gene Family in Rapeseed and Tobacco. **International Journal of Molecular Sciences**, v.18, n.12, p.2768, 2017. https://doi.org/10.3390/ijms18122768

FILIZ, E.; OZYIGIT, I.I.; VATANSEVER, R. Genome-wide identification of galactinol synthase (*GolS*) genes in *Solanum lycopersicum* and *Brachypodium distachyon.* **Computational Biology and Chemistry**, v.58, p.149-157, 2015. https://doi.org/10.1016/j.compbiolchem.2015.07.006

GIBBONS, B.J.; ROACH, P.J.; HURLEY, T.D. Crystal structure of the autocatalytic initiator of glycogen biosynthesis, glycogenin. **Journal of Molecular Biology**, v.319, n.2, p.463-477, 2002. https://doi.org/10.1016/S0022-2836(02)00305-4

GOODSTEIN, D.M.; SHU, S.; HOWSON, R.; NEUPANE, R.; HAYES, R.D.; FAZO, J.; ROKHSAR, D.S. Phytozome: a comparative platform for green plant genomics. **Nucleic Acids Research**, v.40, n.D1, p.1178-1186, 2012. https://doi.org/10.1093/nar/gkr944

HESLOP-HARRISON, J.S.; SCHWARZACHER, T. Domestication, genomics and the future for banana. **Annals of Botany**, 100: 1073-1084, 2007. https://doi.org/10.1093/aob/mcm191

HU, B.; JIN, J.; GUO, A-Y.; ZHANG, H.; LUO, J, GAO. GGSDS 2.0: na upgraded gene feature visualization server. **Bioinformatics**, v.31, n.8, p.1296-1297, 2015. https://doi.org/10.1093/bioinformatics/btu817

JANG, J.H.; SHANG, Y.; KANG, H.K.; KIM, S.Y.; KIM, B.H.; NAM, K.H. Arabidopsis galactinol synthases 1 (*AtGOLS1*) negatively regulates seed germination. **Plant Science**, v.267, p.94-101, 2018. https://doi.org/10.1016/j.plantsci.2017.11.010

KANNAN, U.; SHARMA, R.; KHEDIKAR, Y., GANGOLA, M. P.; GANESHAN, S.; BÂGA, M.; CHIBBAE, R.N. Differential expression of two galactinol synthase isoforms *LcGolS1* and *LcGolS2* in developing lentil (*Lens culinaris* Medik. cv CDC Redberry) seeds. **Plant Physiology and Biochemistry**, v.108, p.422-433, 2016. https://doi.org/10.1016/j.plaphy.2016.08.004

KELLER, F. Transport of stachyose and sucrose by vacuoles of Japanese artichoke (*Stachys sieboldii)* tubers. **Plant Physiology**, v.98, p.442-445, 1992. https://doi.org/10.1104/pp.98.2.442

KUMAR, S.; STECHER, G.; TAMURA, K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. **Molecular Biology and Evolution**, v. 33, n. 7, p. 1870-1874, 2016. https://doi.org/10.1093/molbev/msw054.

LAHUTA, L.B.; PLUSKOTA, W.E.; STELMASZEWSKA, J.; SZABLINSKA, J. Dehydration induces expression of galactinol synthase and raffinose synthase in seedlings of pea (*Pisum sativum* L.). **Journal of Plant Physiology**, v.171, n.14, p.1306-1314, 2014. https://doi.org/10.1016/j.jplph.2014.04.012

LESCOT, T. La diversité génétique des bananiers. **FruiTrop**, v.221, p.98-102, 2014. Avaiable from: <http://agritrop.cirad.fr/573420/> Acessed: 10 abr. 2018.

LIU, Y.; ZHANK, L.; CHEN, L.; MA, H.; RUAN, Y.; XU T.; QI, M. Molecular cloning and expression of an encoding galactinol synthase gene (*AnGolS1*) in seedling of *Ammopiptanthus nanus*. **Scientific Reports**, v.6, n.36113, 2016. https://doi.org/10.1038/srep36113

NISHIZAWA, A.; YABUTA, Y; SHIGEOKA, S. Galactinol and raffinose constitute a novel function to protect plants from oxidative damage. **Plant Physiology**, v.147, p.1251-1263, 2008. https://doi.org/10.1104/pp.108.122465

PERRIER, X.; BAKRY, F.; CARREEL, F.; JENNY, C.; HORRY, J.P.; LEBOT. V.; HIPPOLUTE, I. Combining biological approaches to shed light on the evolution of edible bananas. **Ethnobotany Research Applied**. v.7, p.199-216, 2009. Avaiable from: <http://www.ethnobotanyjournal.org/vol7/i1547-3465-07-199.pdf> Acessed: 10 abr. 2018. https://doi.org/10.17348/era.7.0.199-216

PRICE, N.S. The origin and development of banana and plantain cultivation. **Bananas and Plantains**, World Crop Serie. Springer, Dordrecht, p.1-13, 1995.

SALOMÃO, L.C.C.; SIQUEIRA, D.L. **Cultivo da Bananeira**. Viçosa, MG: Editora UFV, p.109, 2015.

DOS SANTOS, T.B.; BUDZINSKI, I.G.; MARUR, C.J.; PETKOWICZ, C.L.; PEREIRA, L.F.; VIEIRA, L.G. Expression of three *galactinol synthase* isoforms in *Coffea arabica* L. and accumulation of raffinose and stachyose in response to abiotic stresses. **Plant Physiology and Biochemistry**, v.49, n.04, p.441-448, 2011. https://doi.org/10.1016/j.plaphy.2011.01.023

DOS SANTOS, T.B.; LIMA, R.B.D.; NAGASHIMA, G.T.; PETKOWICZ, C.L.D.O.; CARPENTIERI-PÍPOLO, V.; PEREIRA, L.F.P.; VIEIRA, L.G.E. *Galactinol synthase* transcriptional profile in two genotypes of *Coffea canephora* with contrasting tolerance to drought. **Genetics and Molecular Biology**, v.38, n.2, p.182-190, 2015. https://doi.org/10.1590/S1415-475738220140171

SCHNEIDER, T.; KELLER, F. Raffinose in chloroplasts is synthesized in the cytosol and transported across the chloroplast envelope. **Plant and Cell Physiology**, v.50, n.12, p.2174-2182, 2009. https://doi.org/10.1093/pcp/pcp151

SENGUPTA, S.; MUKHERJEE, S.; PARWEEN, S.; MAJUMDER, A.L. Galactinol synthase across evolutionary diverse taxa: Functional preference for higher plants?. **FEBS Letters**, v.586, n.10, p.1488-1496, 2012. https://doi.org/10.1016/j.febslet.2012.04.003

SENGUPTA, S.; MUKHHERJEE, S.; BASAK, P.; MAJUMDER, A.L. Significance of galactinol and raffinose family oligosaccharide synthesis in plants**. Frontiers in Plant Science** v.6, p.656, 2015. https://doi.org/10.3389/fpls.2015.00656

SIMMONDS, N.W.; SHEPERD, K. The taxonomy and origins of the cultivated bananas. **Botanical Journal of the Linnean Society**, v.55, n.359, p.302-312, 1995. https://doi.org/10.1111/j.1095-8339.1955.tb00015.x

SPRENGER, N.; KELLER, F. Allocation of raffinose family oligosaccharides to transport and storage pools in *Ajuga reptans*: the roles of two distinct galactinol synthases. **The Plant Journal**, v.21, p. 249-258, 2000. https://doi.org/10.1046/j.1365-313x.2000.00671.x

TAJI, T.; OHSUMI, C; IUCHI, S.; SEKI, M.; KASUGA, M.; KOBAYASHI, M.; YAMAGUCHI-SHINOZAKI, K.; SHINOZAKI, K. Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in *Arabidopsis thaliana*. **The Plant Journal**, v.29, p.417-42, 2002. https://doi.org/10.1046/j.0960-7412.2001.01227.x

WANG, Z.; ZHU, Y.; WANG, L.; LIU, X.; LIU, Y.; PHILIPS, J.; Deng, X. A WRKY transcription factor participates in dehydration tolerance in *Boea hygrometrica* by binding to the W-box elements of the galactinol synthase (*BhGolS1*) promoter. **Planta**, v.230, n.1155, 2009. https://doi.org/10.1007/s00425-009-1014-3

WANG, D.; YAO, W.; SONG, Y.; LIU, W.; WANG, Z. Molecular characterization and expression of three galactinol synthase genes that confer stress tolerance in *Salvia miltiorrhiza.* **Journal of Plant Physiology**, v.169, n.18, p.1838–1848, 2012. https://doi.org/10.1016/j.jplph.2012.07.015

YOU, J.; WANG, Y.; ZHANG, Y.; DOSSA, K.; LI, D.; ZHOU, R.; ZHANG, X. Genome-wide identification and expression analyses of genes involved in raffinose accumulation in sesame. **Scientific Reports**, v.8, n.4331, 2018. https://doi.org/10.1038/s41598-018-22585-2

ZHOU, Y.; LIU, Y.; WANG, S.; SHI, C.; ZHANG, R.; RAO, J.; WEI, C. Molecular cloning and characterization of galactinol synthases in *Camellia sinensis* with different responses to biotic and abiotic stressors. **Journal of Agricultural and Food Chemistry**, v.65, n.13, p.2751-2759, 2017. https://doi.10.1021/acs.jafc.7b00377.

Recebido para publicação em 22/10/2018
Revisado em 13/11/2018
Aceito em 28/11/2018