*Available online at* www.pelagiaresearchlibrary.com

Pelagia Research Library

European Journal of Experimental Biology, 2013, 3(4):158-164

ISSN: 2248 –9215
CODEN (USA): EJEBAU

# Phylogenetic study of the 3-hydroxy-3-methylglotaryl-Coenzyme A reductase (HMGR) protein in six different family

**Maryam Darabi[1*] and Samin Seddigh[2]**

[1]*Department of Agronomy and Plant Breeding Sciences, College of Aboureihan, University of Tehran, Tehran, Iran*
[2]*Department of Plant Protection, Varamin-Pishva Branch, Islamic Azad University, Varamin, Iran*

---

**ABSTRACT**

*Isoprenoids are synthesized by condensation of two $C_5$ units' isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). The 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR, EC 1.1.1.34) catalyzes the conversion of HMG-CoA to mevalonate, which is the first committed step in the pathway for isoprenoid biosynthesis. In this study, we did bioinformatics analyses on 12 HMGR protein sequences in six different organisms including Arabidopsis thaliana, Drosophila melanogaster, Gossypium hirsutum, Homo sapiens, Mus musculus and Oryza sativa by computational tools to predict the phylogenetic relationship. Sequence comparison analysis revealed there is a high identity between them and phylogenetic analysis indicated that there is a relationship among species of different organisms. According to these results, HMGRs should be derived from a common ancestor.*

**Keywords:** Bioinformatics, HMGR, Mevalonate pathway (MVA), phylogenetic relationship

---

## INTRODUCTION

There are enormous varieties of metabolites in plants that can be classified into two basic groups according to their functions: the primary metabolites and secondary metabolites [1]. The primary metabolites include nucleic acids, proteins, fatty acids and others that participate in plant basic functions. The secondary metabolites influence ecological interactions between plants and environment [2]. Structurally and functionally diverse isoprenoids (also called terpenoids) are among the largest and the most structurally varied groups of natural products with over 30,000 known compounds, many of which have essentially biological functions in plants, i. e., sterols as the essential components of bio-membranes, carotenoids and chlorophyllins as photosynthesis pigments, plant hormones as the regulators of plant growth, development and defense isoprenoids. Additionally, some isoprenoids are important economically chemicals including flavors, pigments, waxes, rubbers, vitamins, taxol [3], artemisinin [4] and ginkgolides [5]. There are two distinct pathways for isoprenoid biosynthesis in plant kingdom: the well-studied mevalonate (MVA) pathway and deoxyxylulose 5-phosphate/2-C-methyl-D-erythritol 4-phosphate (MEP) pathway (also called DXP pathway), which is mevalonate-independent pathway. Furthermore, these two distinct pathways are localized in different subcellular compartments: the MVA pathway predominates in the cytosol and the MEP pathway in the plastid [6]. As shown in Figure 1, the 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR, EC 1.1.1.34) catalyzes the conversion of HMG-CoA to mevalonate, which is the first step in the pathway for isoprenoid biosynthesis in plants.

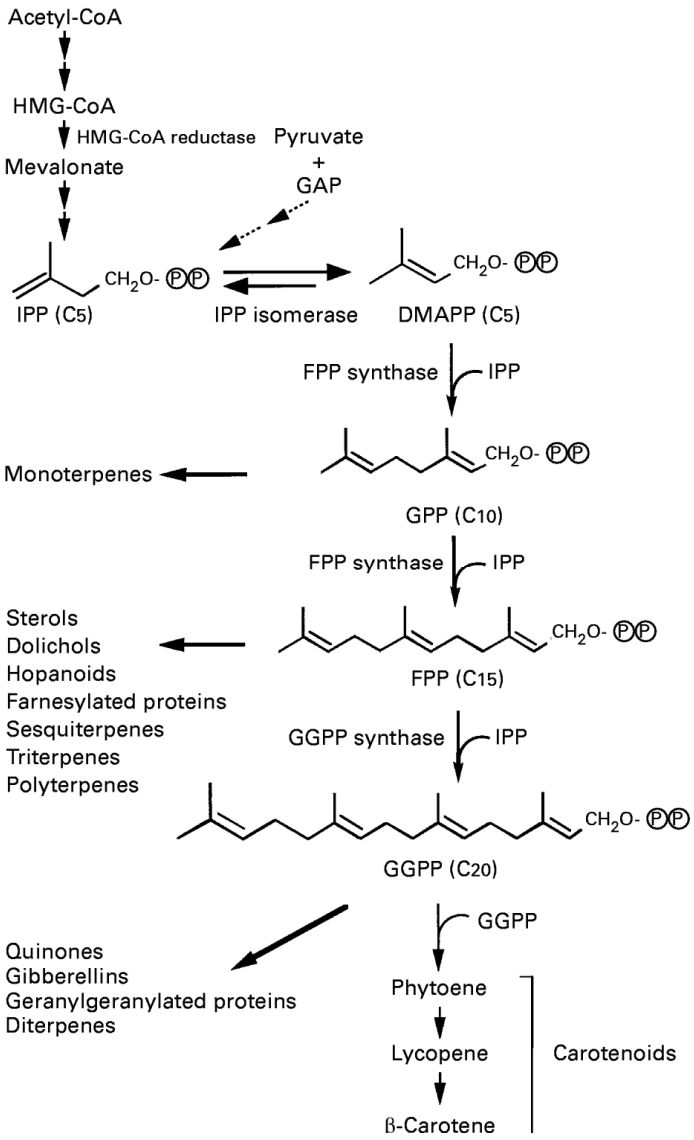



**Figure 1. Schematic representation of mevalonate biosynthetic pathway**
*(Data taken from http://www.biochemj.org)*

HMGR is a rate-limiting enzyme that can be inhibited specifically by lovastatin, catalyzes the NADP-dependent synthesis of mevalonate from HMG-CoA, which is the most important step of the MVA biosynthetic pathway [7]. HMGR is well studied in archeobacteria, bacteria, fungi, plants and animals. In plants, overexpression of 3-hydroxy-3-methylgulutaryl-CoA reductase (HMGR) in the isoprenoid pathway has been attempted to increase triterpenoid productivity [8-12]. Despite improved productivities per unit weight, the transgenic plants sometimes showed growth inhibition, probably caused by metabolic imbalances. In fact, co-overexpression of the HMGR catalytic domain alleviated the growth inhibition caused by the individual overexpression of the HMGR catalytic domain [13].

Structurally, HMGR consists of three regions: the N-terminal region that contains a variable number of transmembrane segments (usually seven in mammals, insects and fungi; two in plants), a linker, and a C-terminal catalytic region of approximately 400 amino-acid residues. Although little sequence similarity is found among the transmembrane domains of HMG-CoA reductases from different species, the C-terminal catalytic region is highly conserved. The structure of this region is predicted to consist of amphipathic helices flanking an extended-pleated sheet [14].

The catalytic regions of HMGR consist of three domains: the small helical amino-terminal N-domain; the large, central L-domain harboring two HMG-CoA binding motifs (EMPIGYVQIP and TTEGCLVA) and a NADP(H)-binding motif (GTVGGGT). Moreover, its architecture resembled a prism with an alpha helix forming the central

structural element; the small helical S-domain harboring a NADP(H)-binding motif characterized by the sequence DAMGMNM [15]. The nucleotides correlated to these motifs are well conserved which can usually be used for designing degenerate primers for isolating the members of *hmgr* gene family. The purpose of this work was applying bioinformatics analysis on HMGR which is a key gene in six different organisms for elucidation of phylogenetic relationship.

## MATERIALS AND METHODS

### Collection of HMGR sequences

All the HMGR protein sequences (FASTA format) belonging to the six different organisms including *Arabidopsis thaliana*, *Drosophila melanogaster*, *Gossypium hirsutum*, *Homo sapiens*, *Mus musculus* and *Oryza sativa* were downloaded from NCBI (http://www.ncbi.nlm.nih.gov; accessed, on June 2013) as listed in Table 1.

**Table 1. The six species analyzed in this study**

| Index | Abbreviation | Species | Family | Number of Sequence |
|---|---|---|---|---|
| 1 | *At* | *Arabidopsis thaliana* | Brassicaceae | 2 |
| 2 | *Dm* | *Drosophila melanogaster* | Drosophilidae | 2 |
| 3 | *Gh* | *Gossypiumhirsutum* | Malvaceae | 2 |
| 4 | *Hs* | *Homo sapiens* | Hominidae | 2 |
| 5 | *Mm* | *Musmusculus* | Muridae | 2 |
| 6 | *Os* | *Oryza sativa* | Gramineae | 2 |

### Bioinformatics analyses

Several online web services and software were used for analyses of HMGRs in six different organisms. The phylogenetic analysis of HMGRs was done with DNASTAR package (MegAlign) using CLUSTAL-W with default parameters. Phylogenetic trees were constructed using MEGA (version 5) from aligned sequences. Five methods were used to construct the trees.

## RESULTS AND DISCUSSION

The biosynthesis pathways of isoprenoid have been discovered and investigated. The encoding enzymes genes involved in both the MVA and MEP pathways have been identified. These may lead to the possibilities for bioengineering the isoprenoid biosynthetic pathways. Various novel strategies are likely to further accelerate the progress of isoprenoid metabolic engineering [16-19]. The results of the sequence alignment showed that HMGRs in six different organisms have a high homology with one another. HMGRs contained two HMG-CoA-binding motifs (ENPVGYVQIP and TTEGCLVA) and two NADP(H)-binding motifs (DAMGMNM and GTVGGGT). In this study, 12 HMGR protein sequences belonging to *Arabidopsis thaliana*, *Drosophila melanogaster*, *Gossypium hirsutum*, *Homo sapiens*, *Mus musculus* and *Oryza sativa* were aligned by CLUSTAL-W. The alignment of protein sequences showed high homology among different organisms. Figure 2 shows the percentage identities among proteins in the multiple alignments of HMGR.

**Percent Identity**

Divergence

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | 68.3 | 36.6 | 36.8 | 72.6 | 67.8 | 39.6 | 39.6 | 59.8 | 40.2 | 41.2 | 68.4 | 1 | AtHMGR-NP_177775 |
| 2 | 37.5 | | 42.2 | 41.8 | 69.9 | 69 | 43.4 | 43.4 | 56.3 | 45.2 | 37.7 | 63.3 | 2 | AtHMGR-NP_179329 |
| 3 | 112.3 | 100.3 | | 99.6 | 40.5 | 38.2 | 46.5 | 46.5 | 60.7 | 46.3 | 27.9 | 42.4 | 3 | DmHMGR-ABY20416 |
| 4 | 111.7 | 99.7 | 0.4 | | 40.7 | 38.4 | 46.6 | 46.6 | 60.7 | 47.2 | 30.5 | 42.5 | 4 | DmHMGR-NP_001163703 |
| 5 | 31.2 | 37.2 | 99.1 | 98.6 | | 85.3 | 45.6 | 45.6 | 59.4 | 45.8 | 42.9 | 67.7 | 5 | GhHMGR-AAC05088 |
| 6 | 29.8 | 37.4 | 101.6 | 101.1 | 16.4 | | 43.2 | 43.2 | 61.6 | 43.2 | 41.9 | 67.2 | 6 | GhHMGR-AAC05089 |
| 7 | 103.2 | 90.2 | 82.5 | 82.5 | 89 | 95.9 | | 100 | 95.1 | 93.5 | 28.9 | 43.4 | 7 | HsHMGR-AAG21343 |
| 8 | 103.2 | 90.2 | 82.5 | 82.5 | 89 | 95.9 | 0 | | 95.1 | 93.5 | 28.9 | 43.4 | 8 | HsHMGR-NP_000850 |
| 9 | 55.8 | 60.6 | 55.1 | 55.1 | 52.5 | 48.1 | 5.1 | 5.1 | | 100 | 5.4 | 58.5 | 9 | MmHMGR-AAA37819 |
| 10 | 100.6 | 89.1 | 81.4 | 81.1 | 87.5 | 94.9 | 6.7 | 6.7 | 0 | | 31.2 | 44.3 | 10 | MmHMGR-NP_032281 |
| 11 | 87.6 | 93.5 | 136.5 | 136.5 | 89.5 | 88.5 | 158.1 | 158.1 | 169.5 | 152.2 | | 45.8 | 11 | OsHMGR-NP_001047915 |
| 12 | 39.5 | 47 | 99.2 | 98.6 | 41.2 | 41.8 | 94.2 | 94.2 | 55.2 | 92 | 79.9 | | 12 | OsHMGR-NP_001063541 |
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | | |

**Figure 2. The degree of percentage sequence identity of residues across the six different organisms' protein sequences that are shown in Figure 3**

As shown in Figure 3, motifs 2 (TTEGCLVA), 3 (DAMGMNM) and 4 (GTVGGGT) were more conserved than motif 1 (ENPVGYVQIP) in six different samples.

Phylogenetic trees were constructed based on the amino acid sequences of HMGRs from six different organisms to investigate the evolutionary relationships among them. According to the result of phylogenetic trees, 12 protein sequences of HMGR were defined approximated six groups. The results revealed that HMGRs were derived from an ancestor gene and evolved into different groups and have relationship with one another (Figure 4).

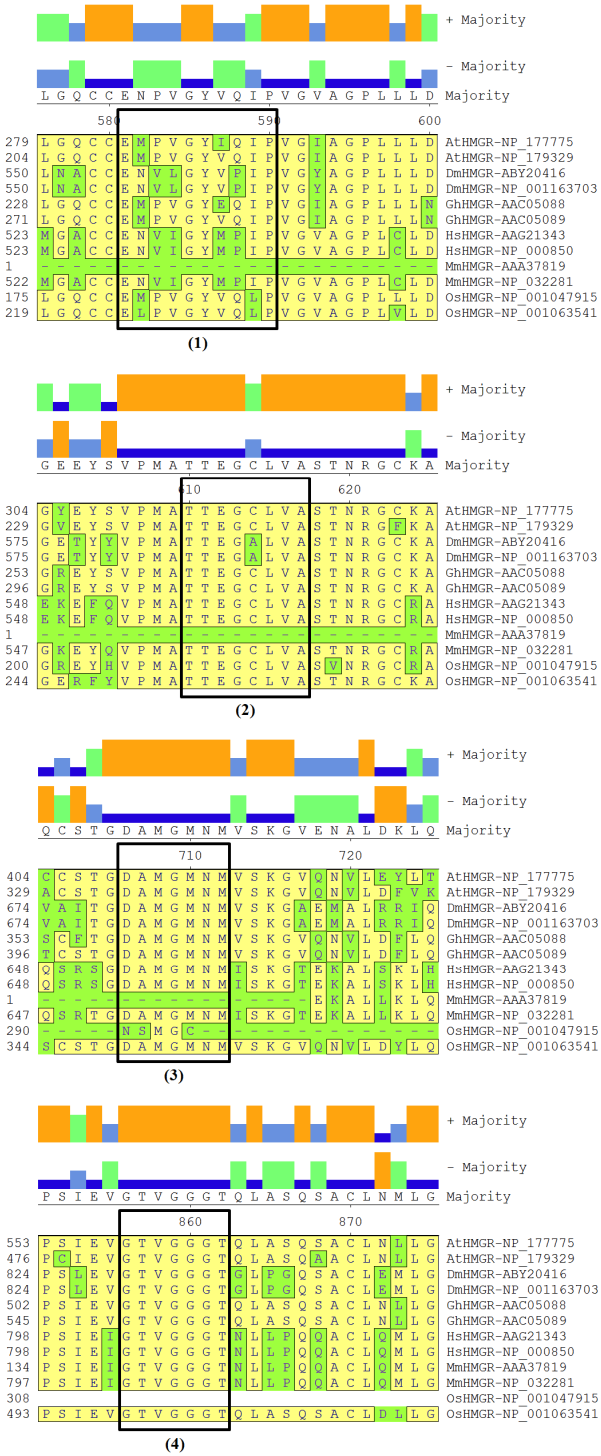
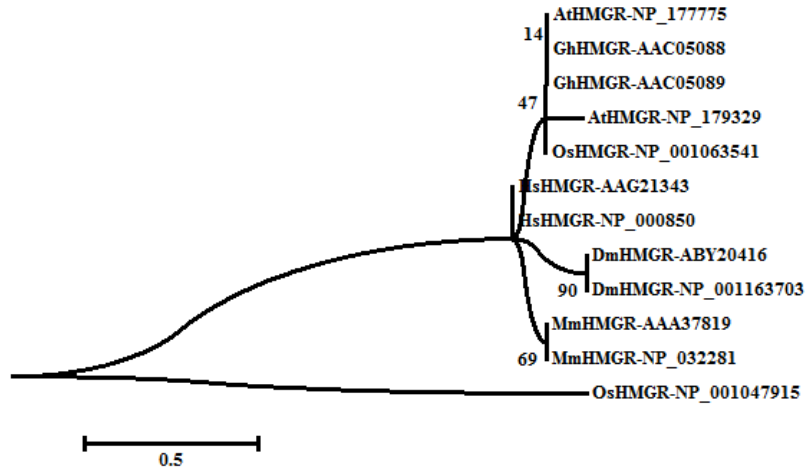
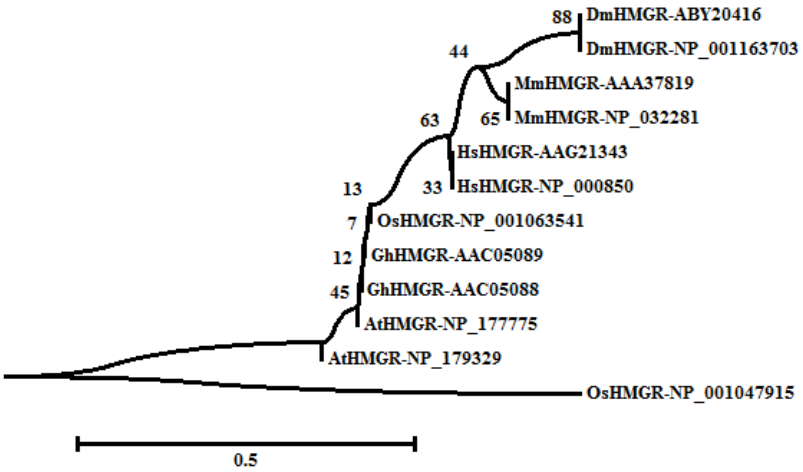


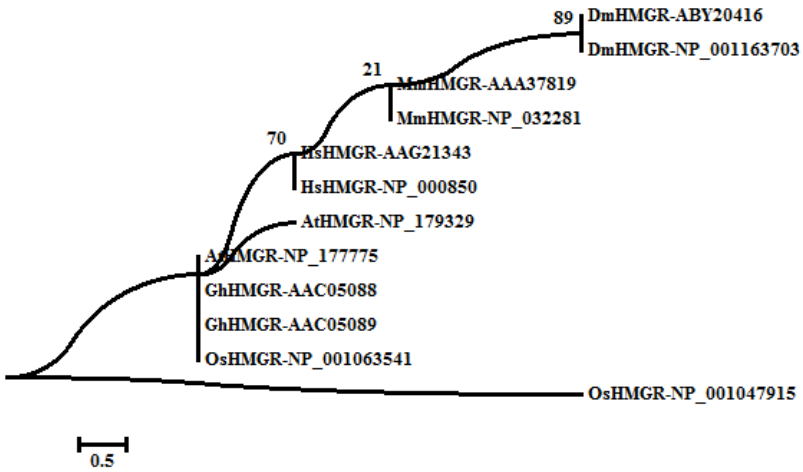

**Figure 3. Multiple alignment of amino acid sequences in six different organisms. Shade with light yellow residues: match the consensus exactly; shade with bright lime residues: differ from the consensus. The putative HMG-CoA-binding sites (1, 2, 3 and 4) were indicated in the shape. Alignments were performed with CLUSTAL-W method**




**Maximum Likelihood Tree**




**Minimum Evolution Tree**




**Maximum Parsimony Tree**

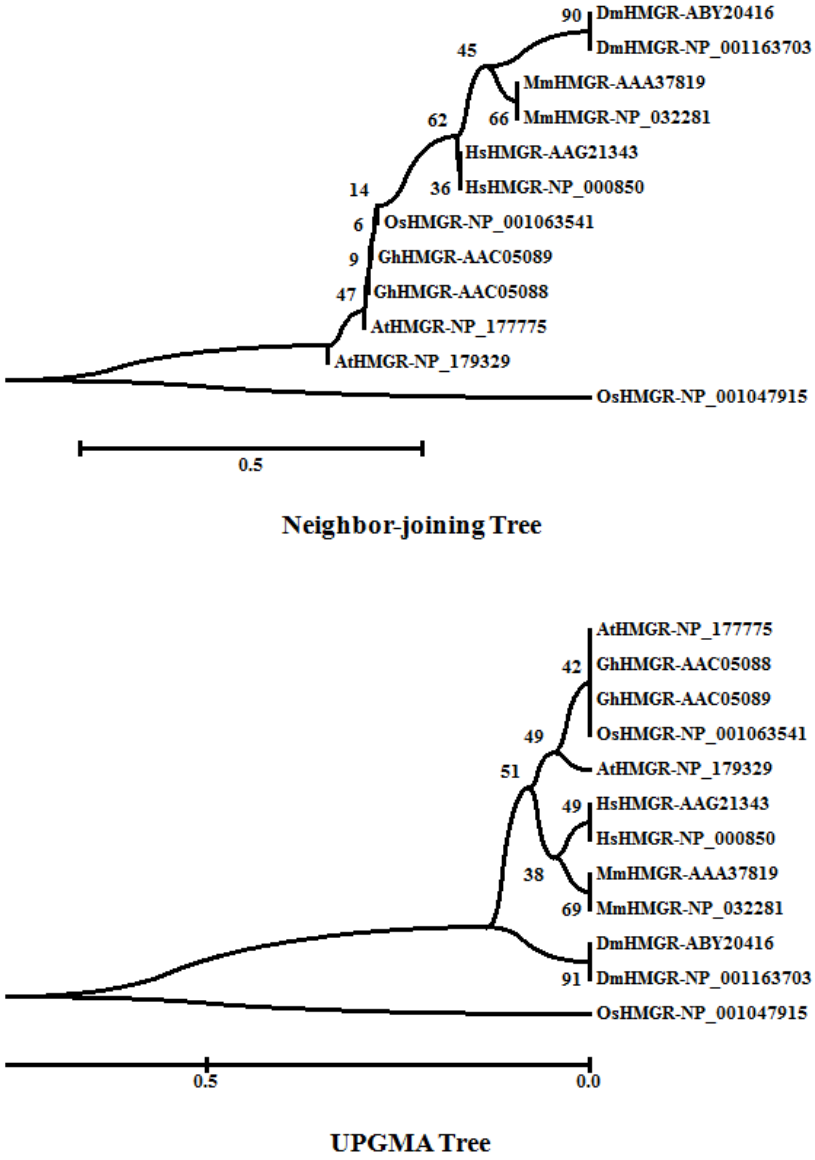



**Figure 4. The phylogenetic trees of HMGRs from six different organisms using the CLUSTAL-W (MEGA 5) program. The Maximum Likelihood, Minimum Evolution, Maximum Parsimony, Neighbor-Joining and UPGMA methods were used to construct the trees. The percentage of 1000 bootstrap replicates was given at each node**

Recently, bioinformatics and genomic tools have revolutionized the studies of organisms metabolism [20, 21]. The novel technologies, including engineering the isoprenoid-based plant defenses, improving the quality of crops, such as “Golden Rice” [16], breaking the isoprenoid biosynthetic bottleneck in plants to produce a higher amount of essential oils in peppermint [22], rebuilding a pathway in bacteria to produce pharmaceuticals such as artemisinin [23], altering the metabolic flux to the desirable direction to yield high-content pharmaceutical agents such as terpenoidindole alkaloids [24], blocking the checkpoint along the pathways to develop new herbicides [25] and anti-parasite drugs [26]. Furthermore, isoprenoid biosynthesis is also important for plant growth and biomass properties [27]. Even though good progress has been made in the metabolic engineering of isoprenoids, it is still a largely undiscovered field [28, 29].

## CONCLUSION

To identify phylogenetic relationships among six different organisms including *Arabidopsis thaliana*, *Drosophila melanogaster*, *Gossypium hirsutum*, *Homo sapiens*, *Mus musculus* and *Oryza sativa*, bioinformatics analyses on 12 HMGR protein sequences were done. Different computational tools were used in this study. The alignment was performed by MegAlign and the phylogenetic trees were constructed by Molecular Evolutionary Genetics Analysis (MEGA), version 5. Sequence comparison analysis showed that there is a high identity between these six species and phylogenetic analysis indicated that there is a relationship among species of different organisms. These results

of phylogenetic tree showed HMGR from different creatures have a relationship with each other, also it can be proved they were derived from an ancestor gene and evolved into different groups.

**Abbreviations**

DMAPP, dimethylallyl diphosphate; DXP, 1-deoxy-D-xylulose 5-phosphate; HMGR, 3-hydroxy-3-methylglutaryl coenzyme A reductase; IPP, isopentenyl diphosphate; MEP, 2C-methyl-D-erythritol 4-phosphate; MVA, mevalonic acid

## REFERENCES

[1] T. Hartmann, *Entomol. Gen Application*, **1996**, 80, 177.
[2] R. Croteau, T. M. Kutchan, N. G. Lewis, In: B. Buchanan, W. Gruissem, R. Jones (Eds Natural products (secondary metabolites) (Am. Soc. Plant Biologists, Rockville, MD, USA, **2000**) 24.
[3] M. C. Wani, H. L. Taylor, M. E. Wall, P. Coggon, A. T. McPhail, *J. Am. Chem. Soc*. **1971**, 93, 2325.
[4] X. Y. Chen, Z. H. Xu, *Med. Chem. Res.*, **1996**, 6, 215.
[5] A. Cieza, P. Maier, E. Poppel, *Fortschritte Der Medzin*, **2003**, 1, 10.
[6] J. A. Bick, B. M. Lange, *Arch. Biochem. Biophys.*, **2003**, 415, 146.
[7] C. Caelles, A. Ferrer, L. Balcells, F. G. Hegardt, A. Boronat, *Plant Mol. Biol.*, **1989**, 13, 627.
[8] J. Chappell, F. Wolf, J. Proulx, R. Cuellar, C. Saunders, *Plant Physiol.*, **1995**, 109, 1337.
[9] H. Schaller, B. Grausem, P. Benveniste, M. L. Chye, C. T. Tan, Y. H. Song, N. H. Chua, *Plant Physiol.*, **1995**, 109, 761.
[10] M. Harker, N. Holmberg, J. C. Clayton, C. L. Gibbard, A. D. Wallace, S. Rawlins, S. A. Hellyer, A. Lanot, R. Safford, *Plant Biotechnol. J.*, **2003**, 1, 113.
[11] S. J. Hey, S. J. Powers, M. H. Beale, N. D. Hawkins, J. L. Ward, N. G. Halford, *Plant Biotechnol. J.*, **2006**, 4, 219.
[12] J. Muñoz-Bertomeu, E. Sales, R. Ros, I. Arrillaga, J. Segura, *Plant. Biotechnol. J.*, **2007**, 5, 746.
[13] D. Manzano, X. Fernández-Busquets, H. Schaller, V. González, A. Boronat, M. Arró, A. Ferrer, *Planta*, **2004**, 219, 982.
[14] E. S. Istvan, J. Deisenhofer, *Biochem. Biophys. Acta*, **2000**, 1529, 9.
[15] J. Ruiz-Albert, E. Cerda-Olmedo, L. M. Corrochano, *Mol. Genet. Genomics*, **2001**, 266, 768.
[16] X. D. Ye, S. A. Babili, A. Klöti, J. Zhang, P. Lucca, P. Beyer, I. Potrykus, *Science*, **2000**, 287, 303.
[17] S. Romer, P. M. Bramley, P. D. Fraser, J. W. Kiano, C. A. Shipton, N. Misawa, W. Schuch, *Nat. Biotechnol.*, **2000**, 18, 666.
[18] V. Mann, M. Harker, I. Pecker, J. Hirschberg, *Nat. Biotechnol.*, **2000**, 18, 888.
[19] C. Schmidt-Dannert, D. Umeno, F. H. Arnold, *Nat. Biotechnol.*, **2000**, 18, 750.
[20] M. Darabi, A. Masoudi-Nejad, G. Nemat-Zadeh, *Mol. Biol. Rep.*, **2012**, 39, 8925.
[21] M. Darabi, S. Seddigh, *Annals of Biological Research*, **2013**, 4(4), 158.
[22] B. M. Lange, R. E. B. Ketchum, R. Croteau, *Plant Physiol.*, **2001**, 127, 305.
[23] V. J. J. Martin, D. J. Pitera, S. T. Withers, J. D. Newman, J. D. Keasling, *Nat. Biotech.*, **2003**, 21, 796.
[24] L. V. D. Fits, J. Memelink, *Science*, **2000**, 289, 295.
[25] B. M. Lange, R. Croteau, *Curr. Opin. Plant Biol.*, **1999**, 2, 139.
[26] H. Jomaa, J. Wiesner, S. Sanderbrand, B. Altincicek, K. Weidemeyer, M. Hintz, I. Turbachova, M. Eberl, J. Zeidler, H. K. Lichtenthaler, D. Soldati, E. Beck, *Science*, **1999**, 285, 1573.
[27] M. E. Eriksson, M. Israelsson, O. Olsson, T. Mortiz, *Nat. Biotechnol.*, **2000**, 18, 784.
[28] A. Soriente, M. De Rosa, A. Scettri, G. Sodano, M. C. Terenico, M. Payá, M. J. Alcaraz, *Curr. Med. Chem.*, **1999**, 6, 415.
[29] P. D. Matthews, E. T. Wurtzel, *Appl. Microbiol. Biotechnol.*, **2000**, 53, 396.