Environment and Ecology 36 (4A) : 1125—1132, October—December 2018
Website: environmentandecology.com ISSN 0970-0420

Cloning and *In silico* Characterization of *Phosphomevalonate kinase* Gene from *Panax sokpayensis* Shiva K Sharma & Pandit

Bhusan Gurung, Dipanwita Saha, Pardeep K. Bhardwaj, Dinabandhu Sahoo

Received 29 May 2018 ; Accepted 2 July 2017; Published on 21 July 2018

Abstract *Panax sokpayensis* Shiva K Sharma & Pandit is an important ginsenoside producing medicinal plant found in Sikkim Himalaya, India. The ginsenosides are produced through ginsenoside biosynthetic pathway whose starting substrates are produced by mevalonate (MVA) and methylerythritol phosphate pathways. *Phosphomevalonate kinase* is an important enzyme which catalyzes the conversion of mevalonate-5 phosphate to mevalonate-5 diphosphate in MVA pathway. In the current study, a cDNA encoding *Phosphomevalonate kinase* was cloned from *P. sokpayensis (PsPMVK)* using degenerate primers and rapid amplification of cDNA ends (RACE) approach. The full length cDNA of *PsPMVK* contained a total of 1806 bp with open reading frame of 1518 bp encoding 505 amino acids and 288 bp 3′ untranslated region including a poly A tail. Bioinformatics analysis revealed that deduced *PsPMVK* protein has predicted molecular weight of 54.72 kDa with theoretical pl of 5.39 and hydrophobic in nature. *In silico* analyses of *PsPMVK* revealed that protein contained ERG8 superfamily conserved domain with same variation at C-terminal region. Phylogenetic analysis showed that *PsPMVK* has closeness to *PMVKs* from other *Panax* species. The predicted secondary structure analysis showed the presence of alpha helix (42.57%), beta turns (7.92%), random coils (30.89%) and extended strands (18.61%). These results provide a foundation for further investigating *PsPMVK* gene functions to regulate the ginsenoside biosynthesis in *P. sokpayensis*.

Keywords Mevalonate pathway, *Phosphomevalonate kinase,* RACE, *Panax sokpayensis.*

Introduction

In medicinal plants, it has been well established that various secondary metabolites are synthesized during normal growth and development. Methylerythritol phosphate (MEP) pathway and the mevalonate/mevalonic acid (MVA) pathway are responsible for producing precursors for the biosynthesis of terpenoids, the largest group of secondary metabolites in plants. In *Panax* spp., MVA pathway along with the MEP pathway produces precursors for the biosynthesis of triterpenoid saponins called ginsenosides (Zhao et al. 2014). The MVA pathway converts acetyl CoA through six sequential steps to isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), the precursors for the terpenoid biosynthesis.

Bhusan Gurung[1], Pardeep K. Bhardwaj[1]*, Dinabandhu Sahoo[1]
[1]Institute of Bioresources and Sustainable Development, Sikkim Center, Tadong, Gangtok 737102, Sikkim, India

Dipanwita Saha[2]
[2] Department of Biotechnology, University of North Bengal, Siliguri 734013, Darjeeling, India
e-mail: pardeepbiotech@rediffmail.com
*Corresponding author

Phosphomevalonate kinase (*PMVK*) catalyzes the fifth reaction of the pathway in which mevalonate-5-phosphate is phosphorylated to mevalonate-5-diphosphate in the presence of Mg^{2+} as co-factor (Miziorho 2011, Olivier et al. 1999). In plants, cDNAs belonging to PMVKs have been cloned and characterized from *Matricaria chamomilla* (Xu et al. 2018), *Catharantus roseus* (Simkin et al. 2011), *Hevea brasiliensis* (Sando et al. 2008) and *Chamaemelum nobile* (Yan et al. 2016).

Panax sokpayensis Shiva K Sharma & Pandit is a perennial medicinal herb found in Sikkim Himalaya (Sharma and Pandit 2009). Like other *Panas* spp. that are traded extensively, *P. sokpayensis* also contains triterpenoid saponins, known as ginsenosides which are responsible for its medicinal properties (Gurung et al. 2018). The rhizome of this herb is used in the preparation of herbal medicines by the traditional healers and is also traded commercially (Badola and Pradhan 2013, Gurung et al. 2016). These ginsenosides have cardioprotective, immunomodulatory, antifatigue, anticancerous, antidiabetic and antioxidant properties (Lee and Kim 2014, Xiao et al. 2015). However, it is difficult to obtain sufficient quantity of these metabolites from natural resources. Moreover, chemical synthesis of these triterpenoid saponins would be tasking and very expensive. Therefore, detailed understanding of enzymes involved in triterpenoids biosynthesis and their activities in *Panax* spp. is required for improving the production of these metabolites.

Phosphomevalonate kinase is a phosphotransferase which catalyzes the conversion of mevalonate-5-phosphate to mevalonate-5-diphosphate and is identified as regulatory enzyme of mevalonate pathway (Garcia and Keasling 2014). Full length sequences of different plant *PMVKs* have been submitted in NCBI Genbank. For example, many *PMVKs* have been reported from transcriptome studies of various *Panax* spp. (Li et al. 2013, Sun et al. 2010, Wang et al. 2016). Considering the important role of *PMVK* in mevalonate pathway for regulating substrate flux, further characterization is required for its implications in pathway engineering. The present study reports the degenerate primer and rapid amplification of cDNA ends (RACE) based cloning of full length cDNA sequence of *Panax sokapyensis PMVK (PsPMVK)* and its *in silico* analyses to predict its putative polypeptide and various conserved domains.

(Authors acknowledge the Department of Biotechnology (DBT), Government of India for providing funds through IBSD core grant. Department of Forest, Environment and Wildlife Management, Government of Sikkim is also acknowledged for facilitating the research permit to collect plant from Sikkim Himalaya. BG acknowledge DBT for providing financial assistance).

Materials and Methods

Plant material, RNA isolation and cDNA synthesis

Plants of *P. sokpayensis* were collected with intact rhizomes from Sopakha (altitude 2200 m; N 27^{0}16′57′′, E 088^0 05′ 05′′), West Sikkim, India and maintained in net house at IBSD, Sikkim center, (altitude 1230 m; N 27^0 18′ 41′′, E 088^0 35′ 44′′), Tadong, Gangtok, East Sikkim as described previously (Gurung et al 2016). Rhizomes of 5-years old plants were harvested in liquid nitrogen and stored at – 80^{0}C until required. Total RNA was isolated from rhizome following the protocol described by Ghawana et al. (2011). Removal of DNA contamination and first strand cDNA synthesis was carried out using Superscript® III first strand synthesis kit (Invitrogen, USA) as described previously (Gurung et al. 2016).

Degenerate primers based amplification of genes

The degenerate primers (forward : 5′ ATGGCTGTWGTTGCTTCTGC-3′, reverse: 5′-CCATTAMAGGAAGTCATTKC-3′) were designed based on the corresponding gene sequences reported from other plants. PCR reaction was performed using cDNA (1 ul), forward and reverse primers (200 nM final concentrations) and Taq polymerase (1 unit for 25 µl reaction). The PCR conditions for amplifying partial cDNA of *PsPMVK* are as follows: primary PCR– 94^{0}C, (94^{0}C-30 s, 46 ^{0}C-40 s, 72 ^{0}C-1 min) for 35 cycles and final extention at 72 ^{0}C for 7 min; secondary PCR – (94^{0}C-30s, 46^{0}C-40 s, 72^{0}C-1 min) for 40 cycles and

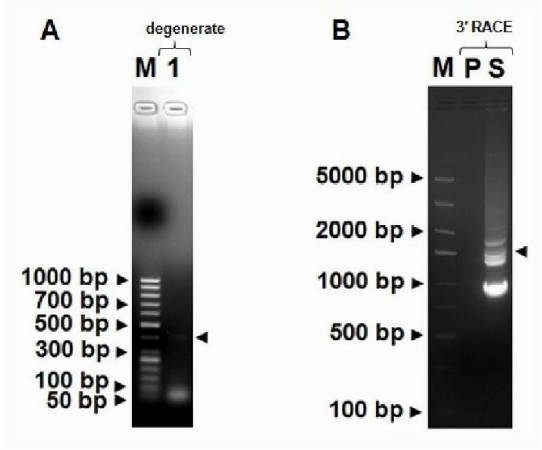

Fig. 1. Agarose gel showing (A) Amplicon obtained using degenerate primers and (B) 3´RACE products of *PsPMVK*. M: DNA marker; P : primary RACE PCR; S : secondary RACE PCR; bp–base pairs; arrow on the right hand side of the gel indicates the desired hand.

final extension at 72 ^{0}C for 7 min. The PCR products were separated in 1.2% agarose gel and the desired amplicon was eluted using gel extraction kit (Sigma-Aldrich Inc., USA).

Cloning and sequencing of desired amplicons

The amplicon was cloned using ta cloning kit (Invitrogen, USA) as per the instructions given in the kit. The presence of the insert was confirmed with a colony PCR with M13 universal primer pair (forward : 5´-GTAAAACGACGGCCAGTG-3´; reverse : 5´-GGAAACAGCTATGACCATG-3´). The PCR was conducted as follows : 94 ^{0}C – 0 min, (94^{0}C-30 s, 52^{0}C-01 min, 72^{0}C-02 min) for 25 cycles. Sequencing was done using the BigDye terminator (version 3.1) cycle sequencing mix (Applied Biosystems, USA) on an automated DNA sequencer (3130_{xl}, Genetic Analyzer, Applied Biosystems, USA) with M13 forward and reverse primers.

Rapid amplification of cDNA ends (RACE)

Based on the sequence of the amplicon obtained through degenerate primer based PCR, the primary and nested primers for 3´ RACE were designed (primary RACE: 5´-TCAAGTCACTCAAGGAACACCTTTGTG-3´, nested RACE:5´-GGCATATGCAAGCCTCGACAATGATAA-3´'). The 3' RACE ready cDNA was prepared using SMARTer^{TM} RACE eDNA Amplification Kit (Clontech, USA) following the manufacturer's protocol. The 3´ primary and nested RACE were conducted using the following thermocycling parameters : 30 cycles of 94 ^{0}C – 10 sec, 68 ^{0}C – 30 sec, 72 ^{0}C – 2 min followed by final extension of 7 min at 72 ^{0}C. The RACE products were run on 1% agarose gel and then cloned and sequenced as mentioned under section 2.3.

In silico analysis of *PsPMVK*

NCBI BLAST algorithm was used for homology search of the full-length *PsPMVK* (www.ncbi.nlm.nih.gov). Putative poly adenylation [poly (A)] signal (PAS) was detected using DNA functional site (DNAFS) miner (http://dnafsminer.bic.nus.edu.sg). The deduced amino acid sequence of *PsPMVK,* its molecular weight and isoelectric point (pl) were predicted using ExPASy proteomics tool (www.expasy.org). Self-Optimized Prediction method With Alignment (SOPMA) was used for secondary structure prediction of the deduced polypeptide sequence (Geourjon and Deleage 1995). Protein plot analysis tool (www.justbio.com) (Kyte and Doolittle 1982) was used to calculate hydropathicity. Multiple sequence alignment was done using ClustalW with default settings (www.genome.jp). The conserved domains were searched against NCBI conserved domain database (Marchler-Bauer and Bryant 2004). The transmembrane regions were searched using TMHMM Server v. 2.0 (http://www.cbs.dtu.dk). Neighbor-joining method was used to construct phylogenetic tree with a bootstrap value of 1000 in MEGA 7 (Kumar et al. 2016) using protein sequences of PMVKs of different organisms retrieved from NCBI GenBank database.

Results and Discussion

PsPMVK amplicon obtained using degenerate primers

A 402 bp amplicon (accession number: KY513110) was obtained using degenerate primer pair (Fig. 1A). On homology search using BLASTn and BLASTx tools at NCBI, the amplicon showed very high degrees of

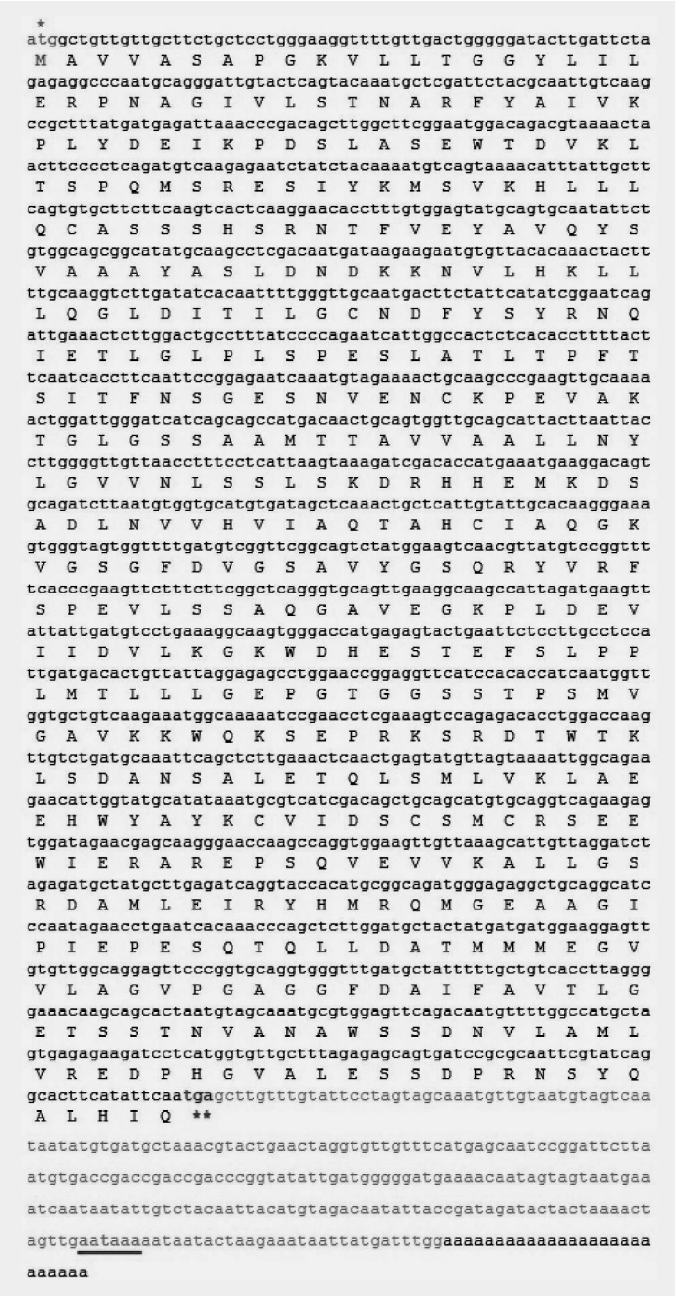

Fig. 2. Nucleotide and deduced amino acid sequence of *P. sokpayensis, Phosphomevalonate kinase (PsPMVK*, MF682468. The polypeptide sequence is represented by a single letter amino acid code under respective codon. The start codon and its corresponding amino acid are colored green and indicated by "*". The stop codon is colored red and indicated by "**". 3′ UTR is colored in blue. PolyA tail is represented by a stretch of adenine residues at the end. Putative poly A signal is bold and underlined by a red bar.

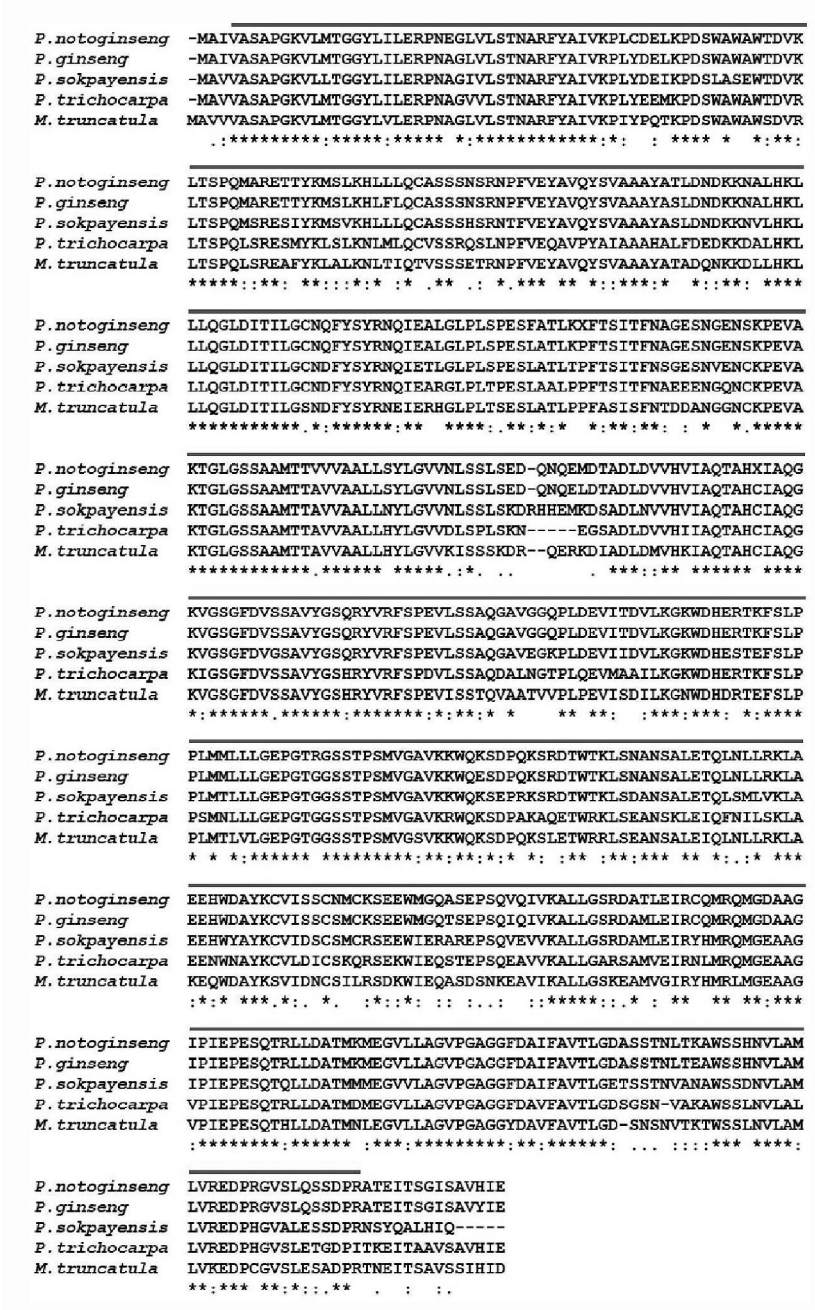

Fig. 3. Multiple sequence alignment of deduced amino acid sequences of *PMVKs* from the *Panax s*pecies and other plants. *Phosphomevalonate kinase* domain is shown by a line in blue. The sequences were retrieved from GenBank.Their accession numbers are as follows: *P. sokpayensis* (MF682468), *P. ginseng (AGZ153214), P. notoginseng (AIK21784), M. truncatula (XP_003602220), P. trichocarpa (XP_002303445).*

similarities with the *PMVKs* of *P. notoginseng* and *P. ginseng* (data not shown), Sequence analysis revealed that partial sequence was amplified from the 5′ region (from starting codon, ATG) and required 3′ RACE for full-length cloning of *PsPMVK*.

Construction and *in silico* analysis of full length cDNA of *PsPMVK*

A fragment of ~ 1500 bp was obtained through 3′ RACE (Fig. 1B). *PsPMVK* of 1806 bp was constructed through the alignment of partial fragment obtained using degenerate primers and 3′ RACE fragment. The sequence was submitted to NCBI with accession number (MF682468). This sequence comprised of 1518 bp long ORF and 288 bp long 3′ UTR which included a 26 bp long adenine repeats representing a section of poly A tail (Fig. 2).

A single polyadenylation signal, AATAAA was detected in the 3′ UTR of *PsPMVK* (Fig. 2). *P. ginseng PMVK* (KC439363) in the NCBI database lacked such PAS while *P. notoginseng PMVK* (KJ804170) contained three PASs (data not shown). PASs have role in 3′end cleavage and polyadenylation of pre-mRNA and formation of different isoforms (Proudfoot 2011). *In silico* analysis predicted a molecular weight of 54.72 kDa for a deduced 505 amino acids of *PsPMVK* and theoretical pl of 5.39 that are similar to that of *PMVK* reported from *Ginkgo biloba, H. brasiliensis, M. chamomilla* and *C. nobile* L. (Sando et al. 2008, Song et al. 2018, Xu et al. 2018, Yan et al. 2016). Multiple sequence alignment of *PMVKs* from across the genera and families from the plant kingdom revealed that the major part of this protein is conserved with some variations at the C – terminal end (Fig. 3). Cladogram formed by phylogenetic analysis of *PMVKs* difierent plant species placed Panax *PMVKs* close to one another indicating their close evolutionary relationships (Fig. 4). The conserved domain search annotated *PsPMVK* with *PMVK* belonging to ERG8 superfamily (superfamily accession number: cl26721) (Fig. 5A). The ERG8 superfamily contains *Phosphomevalonate* kinase orthologues belonging to plants, fungi and eubacteria (Houten and Waterham 2001). Animal and invertebrate *Phosphomevalonate kinases* are non orthologous to the ERG8 type and they belong to P-mevalo_kinase

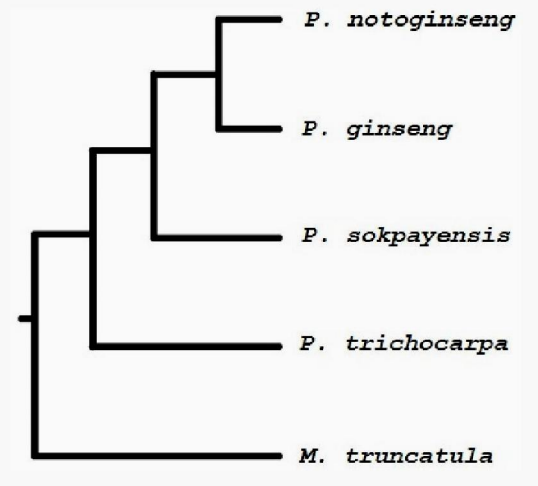

Fig. 4. Phylogenetic tree constructed using UPGMA from the deduced amino acid sequences of *PMVKs* from different plants retrieved from NCBI GenBank. Their accession numbers are as follows: *P. sokpayensis* (MF682468), *P. ginseng (AGZ15314), P. notoginseng (AIK21784), M. truncatula (XP_003602220), P. trichocarpa (XP_002303445).*

superfamily (accession number: cl04466) (Herdendorf and Miziorko 2006, Houten and Waterham 2001). The predicted secondary structure contained 42.57% alpha helix, 18.61% extended strand, 7.92% beta turn and 30.89% random coil (Fig. 5B) is similar to those of *PMVK* from *G. biloba* (Song et al. 2018).

Kyte and Doolittle hydropathy plot predicts two hydrophobic /transmembrane regions (Fig. 5C). However, TMHMM analysis failed to detect any transmembrane helix in *PsPMVK* (Fig. 5D). Though MVA pathway machinery occurs in cytoplasm (Zhao et al. 2014), plant *PMVKs* have been reported to be localized in peroxisomes (Clastre et al. 2011, Simkin et al. 2011). The above findings along with the presence of transmembrane regions in *PsPMVK* tempt us to suggest that *PsPMVK* might be attached to a membrane of some cell organelle.

The transcripts of *PMVK* have also been reported from the transcriptome studies of different *Panax* species (Li et al. 2013, Luo et al. 2011, Sun et al. 2010, Wang et al. 2016, Zhang et al. 2015). *In silico* analysis of *PMVKs* in the of *Panax* spp. have found it to be

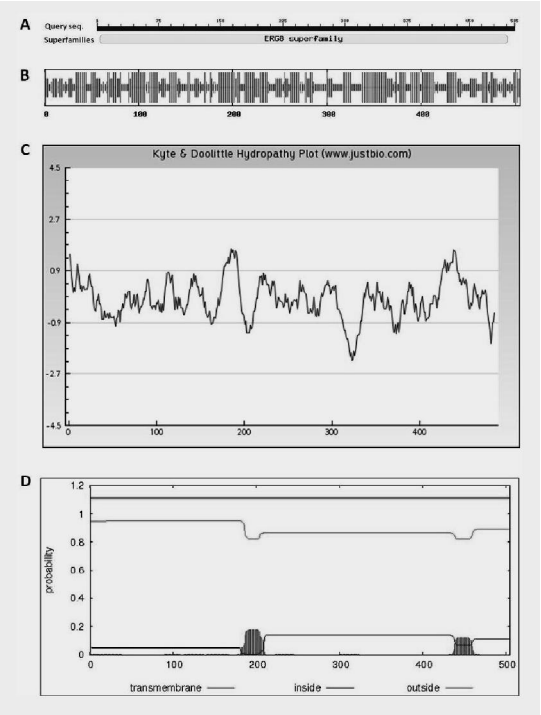

Fig. 5. (A) Prediction of protein identity using NCBI conserved domain analysis (B) Secondary structure prediction of *Phosphomevalonate kinase (PMVK)* using the deduced amino acid sequence. Blue lines indicate alpha helices, purple lines indicate random coils, red lines denote extended strands and green lines represent beta turns (C) Kyte and Doolittle hydropathy plot of *PsPMVK* (D) Prediction of transmembrane region using TMHMM program.

differentially expressed across different developmental stages, among different organs and when treated with methyl jasmonate indicating its probable regulatory role in ginsenoside biosynthesis. Previously, Rat *PMVK*, though non orthologous to plant *PMVKs*, have been found to play a regulatory role in the cholesterol biosynthesis (Olivier et al. 1999). Cloning and characterization of *PsPMVK* will provide a basis for further study the functions of this gene in mevalonate pathway and may help to improve ginsenoside contents by regulating the gene expression. The sequence information can also be used to study 5′ UTR and promoter regions of *PsPMVK* to detect and further characterize the regulatory elements present upstream of the coding sequence.

Conclusion

The current study has successfully cloned cDNA of *P. sokpayensis, Phosphomevalonate kinase* gene belonging to mevalonate pathway using degenerate primer based PCR and RACE. Using bioinformatics tools, putative polyadenylation signal, polypeptide chain, secondary structures and transmembrane regions were predicted. Multiple sequence alignment showed that major portion of the protein was conserved among the plant *PMVKs* while the phylogenetic analysis clearly indicated the close evolutionary relationship among the *Panax, PMVKs*. The genetic information generated in this study can be used for further characterization of this important gene to understand its function and role in regulation of ginsenoside biosynthesis.

References

Badola HK, Pradhan BK (2013) Plants used in healthcare practices by Limboo tribe in South-West of Khangchendzonga Biosphere Reserve, Sikkim, India, Ind J Tradit Knowl 12 : 355—369.

Clastre M, Papon N, Courdavault V, Giglioli-Guivarc'h N, St-Pierre B, Simkin AJ (2011) Subcellular evidence for the involvement of peroxisomes in plant isoprenoid biosynthesis. Pl Signal Behav 6 : 2044—2046.

Garcia DE, Keasling JD (2014) Kinetics of *Phosphomevalonate kinase* from *Saccharomyces cerevisiae*. PLoS One 9 : 1—5.

Geourjon C, Deleage G (1995) SOPMA: Significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Bioinformatics 11 : 681—684.

Ghawana S, Paul A, Kumar H, Kumar A, Singh H, Bhardwaj PK, Rani A, Singh RS, Raizada J, Singh K, Kumar S (2011) An RNA isolation system for plant tissues rich in secondary metabolites. BMC Res Notes 4 : 85.

Gurung B, Bhardwaj PK, Rai AK, Sahoo D (2018) Major ginsenoside contents in rhizomes of *Panax sokpayensis* and *Panax bipinnatifidus*. Nat Prod Res 32 : 234—238.

Gurung B, Bhardwaj PK, Talukdar NC (2016) Subtractive transcriptome analysis of leaf and rhizome reveals differentially expressed transcripts in *Panax sokpayensis*. Funct Integr Genomics 16 : 619—639.

Herdendorf TJ, Miziorko HM (2006) *Phosphomevalonate kinase:*Functional investigation of the recombinant human enzyme. Biochem 45 : 3235—3242.

Houten SM, Waterham HR (2001) Nonorthologous gene displacement of *Phosphomevalonate kinase*. Mol Genet Metab 72 : 273—276.

Kumar S, Stecher G, Tamura K (2016) MEGA 7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33 : 1870—1874.

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157: 105—132.

Lee CH, Kim JH (2014) A review on the medicinal potentials of ginseng and ginsenosides on cardiovascular diseases. J Ginseng Res 38 : 161—166.

Li C, Zhu Y, Guo X, Sun C, Luo H, Song J, Li Y, Wang L, Qian J, Chen S (2013) Transcriptome analysis reveals ginsenosides biosynthetic genes, micro RNAs and simple sequence repeats in *Panax ginseng* C. A. Meyer. BMC Genomics 14 : 245.

Luo H, Sun C, Sun Y, Wu Q, Li Y, Song J, Niu Y, Cheng X, Xu H, Li C, Liu J, Steinmetz A, Chen S (2011) Analysis of the transcriptome of *Panax notoginseng* root uncovers putative triterpene saponin-biosynthetic genes and genetic markers. BMC Genomics 12 : 85.

Marchler-Bauer A, Bryant SH (2004) CD-Search: Protein domain annotations on the fly. Nucleic Acids Res 32 : 327—331.

Miziorko HM (2011) Enzymes of the mevalonate pathway of isoprenoid biosynthesis. Arch Biochem Biophys 505: 131—143.

Olivier LM, Chambliss KL, Gibson KM, Krisans SK (1999) Characterization of *Phosphomevalonate kinase:* chromosomal localization, regulation and subcellular targeting. J Lipid Res 40 : 672—679.

Proudfoot NJ (2011) Ending the message : Poly (A) signals then and now. Genes Dev 25 : 1770—1782.

Sando T, Takao C, Mukai Y, Yamashita A, Hattori M, Ogasawara N, Fukusaki E, Kobayashi A (2008) Clonnig and characterization of mevalonate pathway genes in a natural rubber producing plant, *Hevea brasiliensis*. Biosci Biotechnol Biochem 72 : 2049—2060.

Sharma SK, Pandit MK (2009) A new species of *Panax* L. (Araliaceae) from Sikkim Himalaya, India. Syst Bot 34: 434—438.

Simkin AJ, Guirimand G, Papon N, Courdavault V, Thabet I, Ginis O, Bouzid S, Giglioli-Guivarc'h N, Claster M (2011) Peroxisomal localization of the final steps of the mevalonic acid pathway in planta. Planta 234 : 903—914.

Song Q, Meng X, Liao Y, Zhang W, Ye J, Xu F (2018) Transcriptome-guided gene isolation, characterization and expression analysis of a *Phosphomevalonate kinase* gene (GbPMK) from *Ginkgo biloba*. Int J Agric Biol 20 : 1080—1088.

Sun C, Li Y, Wu Q, Luo H, Sun Y, Song J, Lui EMK, Chen S (2010) De novo sequencing and analysis of the American ginseng root transcriptome using a GS FLX Titanium platform to discover putative genes involved in ginsenoside biosynthesis. BMC Genomics 11 : 262.

Wang J, Li J, Li J, Liu S, Wu X, Li J, Gao W (2016) Transcriptome profiling shows gene regulation patterns inginsenoside pathway in response to methyl jasmonate in *Panax quinquefolium* adventitious root. Sci Rep 6 : 37263.

Xiao D, Yue H, Xiu Y, Sun X, Wang Y, Liu S (2015) Accumulation characteristics and correlation analysis offive ginsenosides with different cultivation ages fromdifferent regions. J Ginseng Res 39 : 338—344.

Xu Y, Liu X, Chang J, Xu F (2018) Cloning and sequence analysis of *Phosphomevalonate kinase* gene *(McPMK)* from *Matricaria chamomilla*. Int J Curr Res Biosci Pl Biol 5 : 10—17.

Yan J, Meng X, Xu F, Chan J (2016) Molecular cloning and sequence analysis of a *Phosphomevalonate kinase* gene (CnPMK) from *Chamaemelum* nobile. Int J Curr Res Biosci Pl Biol 3 : 157—162.

Zhang GH, Ma CH, Zhang JJ, Chen JW, Tang QY, He MH, Xu XZ, Jiang NH, Yang SC (2015) Transcriptome analysis of *Panax vietnamensis* var *fuscidicus* discovers putative ocotillol-type ginsenosides biosynthesis genes and genetic markers. BMC Genomics 16 : 159.

Zhao S, Wang L, Liu L, Liang Y, Sun Y, Wu J (2014) Both the mevalonate and the non-mevalonate pathways are involved in ginsenoside biosynthesis. Pl Cell Rep 33 : 393—400.