Awais et al., Clon Transgen 2018, 7:1
DOI: 10.4172/2168-9849.1000163

Cloning & Transgenesis
ISSN: 2168-9849

Research Article Open Access

Sequence Analysis of HMG-CoA Synthase Like Gene from *Brassica rapa*

Hina Awais[1#], Zahid Mushtaq[1,2*#], Sumaira Sarwar[1] and Amer Jamil[2]

[1]*Bioactive Molecules Research Lab (BMRL), Department of Biochemistry, University of Agriculture, Faisalabad, Pakistan*
[2]*Molecular Biochemistry Lab, Department of Biochemistry, University of Agriculture, Faisalabad, Pakistan*
[#]*Both the Authors Contributed Equally*

Abstract

Terpenoids are synthesized in plants through the mevalonate (MVA) and the methylerythritol phosphate (MEP) pathways, among isopentenyl diphosphate (IPP) as the central intermediate metabolite. 3-Hydroxy-3-methylglutaryl-Coenzyme A synthase (*HMGS*) is a subsequent enzyme in MVA pathway of isoprenoid biosynthesis and catalyses the condensation of acetyl-CoA by acetoacetyl-CoA to yield HMG-CoA. Different genes of *HMGS* are present in different species of plants and animals for the synthesis of secondary metabolites. The present study was planned to compare and analyse bioinformatically the HMG-CoA gene sequence of *Brassica rapa* with other sequences of different species. The gene sequence have 925 bp coding DNA sequence and was submitted in National Centre for Biotechnology Information (NCBI) with accession Number. MF577077. The alignment of nucleotide and protein sequences was carried out through Geneious. Bioinformatics comparison including structural and functional analysis of HMG-CoA synthase like gene sequence isolated from *Brassica rapa* was carried out by using different bioinformatics tools and softwares. These comparisons gave a better information about properties, genetic and functional diversity of this gene along with amino acid substitutions.

Keywords: HMG-CoA synthase like gene; *Brassica rapa*; Brassiceae; Terpenoids; Bioinformatics; Software; Phylogeny

Introduction

Terpenoids are main diverse classes of secondary metabolites in plants. More than twenty five thousand types of terpenoids have been recognized so far in plants. Terpenoids hold multiple Eco physiological roles e.g. regulating heat tolerance of plants, regulating plant growth and development, attracting pollinators, resisting photooxidative tension, providing direct and indirect plant defense. Terpenoids are synthesized in plants by two pathways, MVA (mevalonate) and MEP (2C-methyl-Derythritol 4-phosphate). As the first catalyzing enzyme in the MVA pathway, *HMGS* plays a significant role in the biosynthesis of terpenoids in plants [1-7].

HMGS is a piece of acyl-condensing enzymes family. Considered mass of this enzyme is 105.8 kDa and it has two identical isozymes. This enzyme is stimulated by a compound of Dithioerythreitol. Energy of activation, pH and temperature is 62.5 J mol–1, 8.5 and 35°C respectively [8-15]. In *HMGS* the foremost step of the reaction is a cysteine residue that performs like a nucleophile. Through acetyl-CoA, the acetylation process of the enzyme is occur which form the condensed coenzyme A [16-30].

HMGS (1) and *HMGS* (2) are two *HMGS* isoforms. *HMGS* (1) is implicated in cholesterol synthesis and *HMGS* (2) has a remarkable role in scheming ketogenesis in mammals [22]. HMG (CoA) synthase enzyme of mitochondria has a most important function in the production of the ketone bodies and into the formation of HMG (CoA). Ketone bodies have many important functions such as it take action as a substitute of energy for glucose in a lot of tissues as in heart, muscle, brain and contain a significant part in the duration of metabolic stress and starving situations. The rate-limiting enzyme in the ketogenic pathway is mitochondrial HMG (CoA) synthase [4].

HMGS genes have been found in *Brassica juncea, Arabidopsis thaliana, hevea brasiliensis, Pinus sylvestris, taxus X media, Salvia miltiorrhiza, Solanum lycopersicum* and *oryza sativa* such as *BjHMGS1, HbHMGR1, HbHMGR4, HbHMGS1, HbHMGR5, HbAACT1 and HbHMGS2. HMGS* genes plays very significant and imperative role in different types of the animals. It may also be encodes the chief enzymes in the biosynthesis of the cholesterol and also the controlling of the ketogenesis. *HMGS* genes may also been investigated in numerous types of the animals like *Mesocricetu sauratus* (golden hamster), *Rattusnor vegicus* (rat), pigs *and Blattella germanica* (German cockroach) [18].

In 2009, Besser and colleagues worked on protein sequence analysis of *H. brasiliensis* HMG(S) as they align the concluded amino acid in order of *H. brasiliensis HMG(S)* with other known HMG(S) sequences, including *A. thaliana* (11), *P. sylvestris* (12), human cytosolic (18), human mitochondria1 (9) and *S. pombe* (20) was done to analyze sequence conservation. The *H. brasiliensis* protein showed (81), (74), (51), (49) and (45) percent identity with *A. thaliana, P. sylvestris*, and human cytosolic, human mitochondrial and *S. pombe* correspondingly. Due to different uses of *HMGS* gene we decided to analyze the gene sequence of HMG-CoA synthase by different bioinformatics software and tools to get knowledge that how much genetic and functional diversity was present.

Materials and Methods

Isolation of HMG-CoA synthase like gene

Leaves of the *Brassica rapa* were obtained from botanical garden of University of Agriculture, Faisalabad. The genomic DNA from *Brassica rapa* was isolated by following Doyle and Doyle protocol (1990). Primer designing was done through database information available on NCBI (www.ncbi.com) for HMG-CoA synthase like gene as shown below:

HMG-CoA F1: 5'- TTATGGTGGAACTGCGGCTT -3'

***Corresponding author:** Zahid Mushtaq, Bioactive Molecule Research Lab (BMRL), Department of Biochemistry, University of Agriculture, Faisalabad-38040, Pakistan, Tel: +92-3226058801; E-mail: zahidbiochemist@yahoo.com; zahidmushtaquaf@uaf.edu.pk

Received April 04, 2018; **Accepted** April 17, 2018; **Published** April 23, 2018

Citation: Awais H, Mushtaq Z, Sarwar S, Jamil A (2018) Sequence Analysis of HMG-CoA Synthase Like Gene from *Brassica rapa*. Clon Transgen 7: 163. doi: 10.4172/2168-9849.1000163

HMG-CoA R1: 5'- AGGGTTCGAAGAAGGGAAGTTT -3'

HMG-CoA synthase like gene was amplified from extracted DNA with gradient polymerase chain reaction using specifically designed primers by providing given thermo-cycling conditions: initial denaturation for 4 minutes at 94°C followed by 32 cycles of repeated denaturation at 94°C for 1 minute, annealing at 48°C to 54°C for 1.45 minutes, Extension at 72°C for 1.45 minutes and final extension of 5 min at 72°C. The isolated gene was confirmed by agarose gel (1% w/v in 1x TAE) electrophoresis.

Sequence submission to GenBank

The gene was sequenced from Penicon Company (http://www.penicon.com). Then sequence was processed for cleaning and trimming by using DNA Baser (4.36.0) and a contig was obtained. After that the sequence of HMG-CoA synthase like gene with 925bp was submitted to GenBank database (NCBI) with accession no. MF577077.

Bioinformatics analysis

Genomic and proteomic analysis was done on the selected gene sequence of HMG-CoA synthase by various bioinformatics tools and softwares.

Multiple sequence alignment: The homology of nucleotide and protein sequence of HMG-CoA synthase like gene with the already reported sequences was checked through BLASTn and BLASTp (NCBI) and related or similar *HMGS* gene sequences were aligned by using geneious.

Phylogenetic analysis: The desktop version of MEGA5 was used for the erection of phylogenetic hierarchy from selected sequences of gene which have significant correspondence with our isolated sequence for better understanding of the origin and evolution of the encoded protein.

Characteristic features of HMG-CoA synthase like gene: Different software and tools were utilized for complete sequence analysis such as ExPASy - ProtParam tool (for physiochemical characteristics), SignalP 4.1 (for signal peptide prediction), and Peptide 2.0 (acidic residues, Basic residues, Hydrophobic residues).

Visualization of proteoforms: Protter, the open-source tool was used to visualize proteoforms and interactive integration of annotated and predicted sequence features [23].

Secondary structure prediction: PSIPRED was used to predict secondary structure elements in HMG-CoA synthase like protein from *Brassica rapa* [5].

3-D Structure prediction

Sequences retrieval to predict 3D structure: HMG-CoA synthase from *Brassica juncea* in the apo-form [25] was used as a template to construct a 3D model for *Brassica rapa* HMG-CoA synthase.

Models building by homology modelling: To predict 3-D structures of HMG-CoA synthase from *Brassica rapa*, homology modeling was used that is the most suitable method for building protein models [2]. CPH Model Server was used to select templates [21]. Modeller v9.16 was used, for template and query alignments [10] using align 2d command and output file in PIR format was used for building five models against each query. Root Mean Squared Deviation (RMSD) for superimposition of HMG-CoA synthase was performed by using UCSF Chimera 1.10 work bench [24].

Results and Discussion

Terpenoids are the key residents of different types of essential oils; these are extensively used as flavouring agents and also for the further segregation of the flavouring constituents [6]. *HMGS* play beneficial role in the formation of terpenoids, so due to reveal these important factors about *HMGS* we analyse and compare the HMG-CoA gene sequence of *Brassica rapa*.

Isolation and sequence submission

To isolate the HMG-CoA synthase like gene from *Brassica rapa*, degenerate oligodeoxy ribonucleotides were designed from *Brassica rapa* cultivar Chiifu-401-42 chromosome A2, *B.rapa* 1.0, whole genome shotgun sequence (Accession Number: NC_024796.1). Online bioinformatics software and tools were used to design the primers for the amplification of HMGCoA synthase like gene. Approximately an amplicon of 1220 bp of HMG-CoA synthase like gene was isolated through PCR (Polymerase Chain Reaction). After sequencing a contig of 1162 bp of HMG-CoA synthase like gene was obtained, then for annotation the sequence was cleaned and joined as follows for CDS region (1-61, 151-309, 379-453, 538-612, 700-756, 837- >925). The final selected sequence has 925 bp which was then submitted into genbank NCBI, BanKIt. After getting the accession No. MF577077, the sequence was preceded for bioinformatics analysis.

Alignment of HMG CoA synthase like gene

Basic Local Alignment Search Tool (BLAST) is the most accepted search algorithm and can accommodate nucleotide or protein sequence(s). BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to identify local regions of similarity and statistical significance of nucleotide query sequence of HMG-CoA synthase. Multiple sequence alignment (https://www.geneious.com) was performed through Geneious [16]. The truncated sequences were deleted and longer sequences were shortened in the multiple sequence alignment to make them all equal in length. We used Clustal W and Megs 5 software for alignment and then combine it in Geneious software for better results as shown in Figure 1.

Phylogenetic analysis of HMG CoA synthase like gene

The evolutionary narration was indirect by using the Neighbour-Joining method [26]. The distances were computed by using the Kimura 2-parameter method [17]. The analysis involved 11 nucleotide sequences of different *HMGS* gene from the plants of various families. The query sequence of HMG-CoA synthase like gene which was isolated from *B. rapa* not fall in the *Brassicaceae* family as shown in Figure 2, because this sequence was isolated through wet lab experiments and have both exonic and intronic parts [27-29].

Furthermore, the evolutionary analyses of protein sequence of HMG-CoA synthase like gene was inferred by using the Maximum Likelihood method which based on the Dayhoff matrix based model [27]. The query protein belonged to the Malvids family of plants. There was only one plant of Chenopodiaceae family that appeared close to Malvids family. Thirteen plants were taken from Fabids family and all fell into the same clade. The analysis involved 34 amino acid sequences. Evolutionary analyses were conducted in MEGA 5 [29,30]. The overall mean distance of nucleotide and protein sequences of *B. rapa* from other plant families was found to be 0.300 and 0.128. In the results of evolutionary protein sequences analysis of HMG-CoA synthase like gene, was revealed that our query protein sequence fall within *Brassicaceae* family as shown in Figure 3, because the mRNA

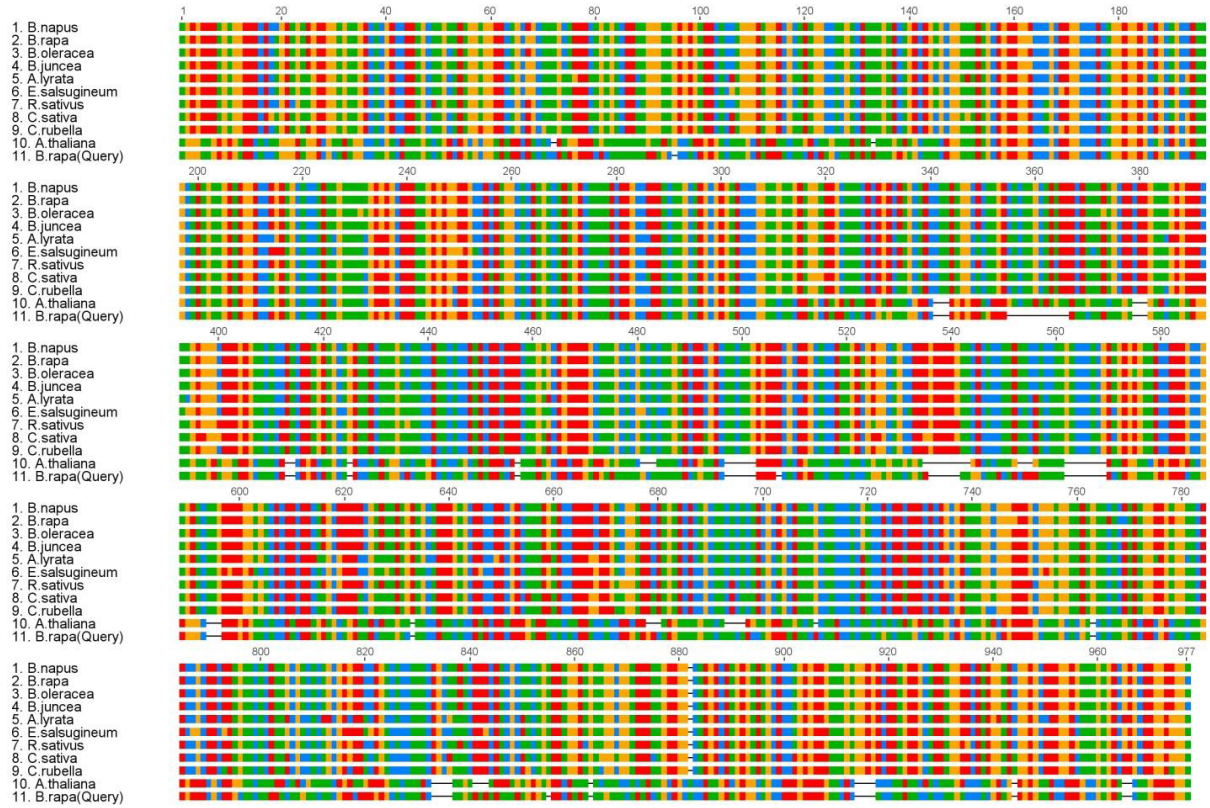

Figure 1: Multiple sequence alignment of nucleotide sequences of HMG-CoA synthase like gene. Similar color bars were showing the aligned sequences.

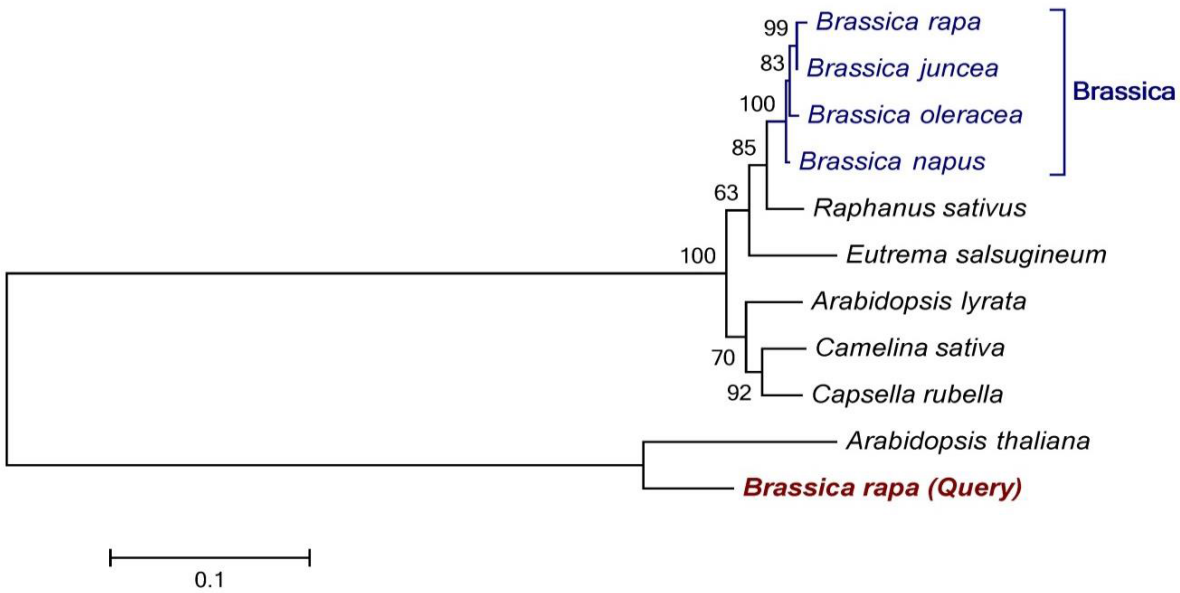

Figure 2: Phylogenetic analysis using neighbor joining method of 11 nucleotide sequences including query sequence of *Brassica rapa*.

sequence of query protein have only exonic part. Irshad et al. in 2017 computing the evolutionary distance of CalY gene by using neighbor-joining, Tamura-Nei method. Strong relationship among CalY protein of different *Bacillus* species was predicted.

DNA Sequence analysis

In bioinformatics, sequence analysis is the process of subjecting a DNA, RNA or peptide sequence to any of a wide range of analytical

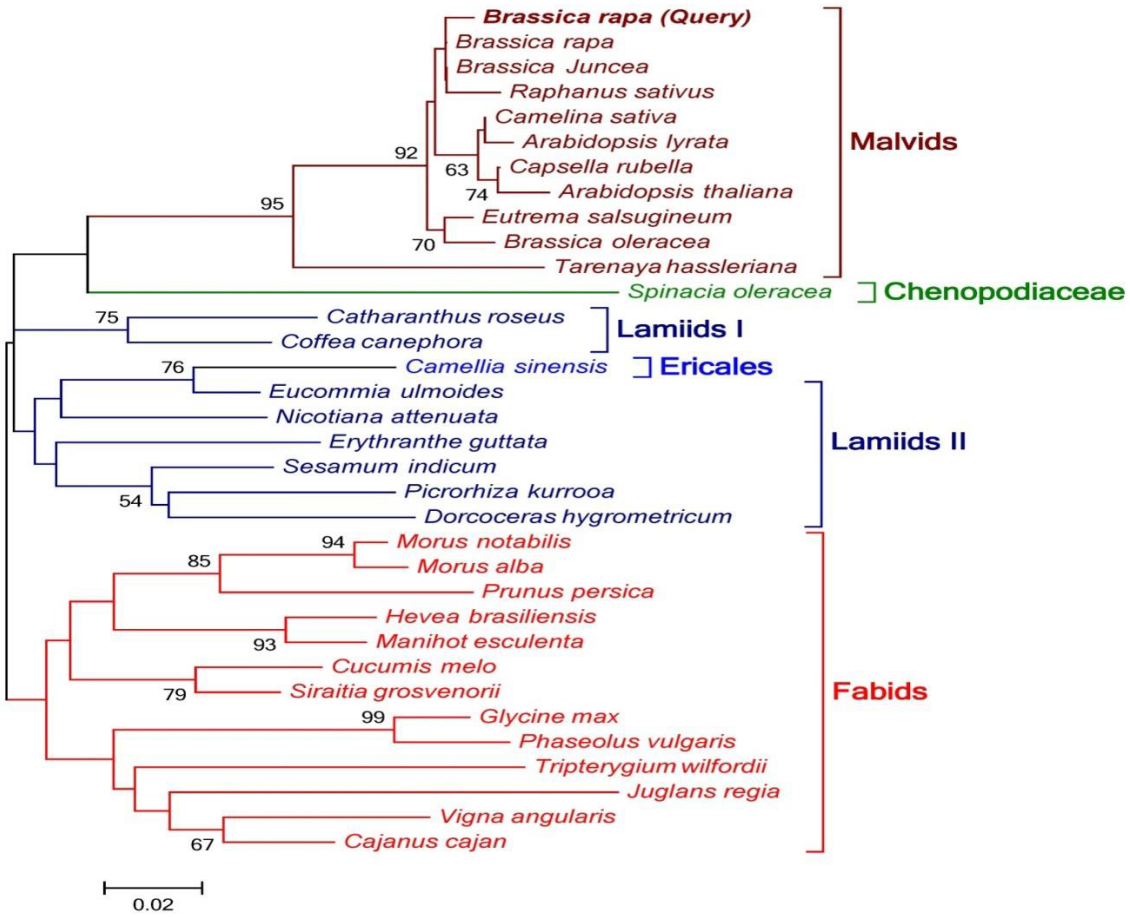

Figure 3: Molecular phylogenetic analysis by maximum likelihood method. 34 protein sequences including query sequence of *Brassica rapa*.

methods to understand its features, function, structure, or evolution. In this analysis we calculate the conserved site, variable site, parsimony site and singleton site by Mega 6. Conserved sites are those sequences that are analogous or the same sequences of nucleic acid(s), protein(s) or in polysaccharide(s) across species or within diverse molecules produced by the same organism paralogous sequences whereas the variable sites contain at least two types of nucleotides or amino acid. Parsimony sites are also called informative sites because it has a position in the set of sequence at which there are at least two different character states are present. A singleton site contains at least two types of nucleotides or amino acids in which one is occurring for multiple times [19]. HMG-CoA synthase like gene have 283 conserved sites, 120 variable sites, 41 parsimony sites and 79 singleton sites. So it is concluded that our gene sequence of *HMGS* has four distinctive sites which contain either the same sequences of biomolecules or the presence of different types of nucleotides and amino acids that gave the information about mutation points in our sequence. Each *HMGS* gene sequence has different number of sites according to its location and source.

Structural and characteristic analysis

The physicochemical properties of a protein are determined by the analogous properties of the amino acids in it. So in proteomic analysis, we evaluated the different parameters of translated sequence of HMG CoA synthase like protein through ExPASy, ProtParam. By the use of this software we predicted physiochemical properties of our gene as shown in Table 1 [11]. Whereas the open reading frame (ORF) of *HMGS* sequence was find out through ORFfinder, NCBI (https://www.ncbi.nlm.nih.gov/orffinder) and the length of ORF1 based on 261 nucleotides and 86 amino-acid residues.

Analysis result was showed that the query protein has 172 residues with iso-electric point 5.18 and was a stable protein because instability index was favourable till 40. Ext. coefficient refers to several different measures of the absorption of light in a protein medium whereas aliphatic index showed that our selected protein was thermo-stable because aliphatic side chain has a positive factor for the increase of thermo-stability of globular protein. GRAVY value for a peptide is calculated as the sum of hydropathy value of all amino acids, divided by the number of residues in the sequence. GRAVY value tells about the solubility of a protein and negative value indicates that HMG-CoA synthase like protein was hydrophilic in nature. Mustafa et al. in 2017 predicted the Physicochemical properties of GDF11 by ProtParam tool and concluded that concentration of a protein in solution can be calculated by using its extinction coefficient as well as the values of aliphatic index, instability index and grand average of hydropathicity (GRAVY) gave idea about the stability of GDF11.

A signal peptide (sometimes referred to as signal sequence, targeting signal, localization signal, localization sequence, transit peptide, leader sequence or leader peptide) is a short peptide (usually 16-30 amino acids long) present at the N-terminus of the majority of

newly synthesized proteins that are destined towards the secretory pathway. Signal-P 4.1 was used for prediction of signal peptides in HMG-CoA synthase-like protein sequence and the result showed that there was no signal peptide in the analysed protein fragment.

The hydrophobicity index is a measure of the relative hydrophobicity, or how soluble an amino acid in water. In a protein, hydrophobic amino acids are likely to be found in the interior, whereas hydrophilic amino acids are likely to be in contact with the aqueous environment. Total percentage of hydrophobicity, acidic, basic and neutral amino acid residues of our desired *HMGS* protein were as following 40.7%, 13.3%, 12.21% and 33.72% which was analysed through peptide 2.0. These scales are commonly used to predict the trans-membrane α helices of membrane proteins. Consecutively measuring amino acids of query protein and changes in values indicate the attraction of specific protein regions towards the hydrophobic region.

A protein domain is a conserved part of a given protein sequence and (tertiary) structure that can evolve, function, and exist independently of the rest of the protein chain. Each domain forms a compact three-dimensional structure and often can be independently stable and folded. Many proteins consist of several structural domains. One domain may appear in a variety of different proteins. Molecular evolution uses domains as building blocks and these may be recombined in different arrangements to create proteins with different functions. Pfam server was used to predict domains in query sequence [9]. Domains of HMG CoA synthase were present on both N and C terminal of chain. There were two results in which first one has envelope results that rely on the probability basis and other were alignment results that represent on the behalf of the sure basis. On N terminal envelope and alignment results were same from 1- 45 residues but on C terminal envelope results vary from 46-172 whereas alignment results from 46-171 residues.

Prediction of protein interactions

STRING is a database of known and predicted protein-protein interactions. The interactions include direct (physical) and indirect (functional) associations; they stem from computational prediction, from knowledge transfer between organisms and from interactions aggregated from other (primary) databases. To predict protein-protein association networks of query sequence with other proteins, STRING was used [28].

Figure 4 showed the interaction of query protein (HMG-CoA synthase) with other proteins of *Brassicacea* family. There are different lines which have the results of curated 15 databases, experimentally determined data, gene neighbourhood, gene fusions, gene co-occurrence and text mining. Small nodes presented that they have low interaction with query protein (*HMGS* protein) and large nodes showed high interaction with query protein. Protein-protein interactions are very important in various biological processes including cellular communication, gene expression, metabolism and immune responses. All proteins which were participated in protein protein interaction belong to thiolase like family. So as further we used InterPro server for protein sequence analysis and classification [1]. The classification of protein binding regions might be useful to decipher protein interaction networks, understand protein functions and design. The analysis revealed that query protein sequence belongs to hydroxyl methylglutaryl-CoA synthase family of proteins and it showed similarities with two domains i.e. hydroxymethylglutaryl-coenzyme A synthase, N-terminal from amino acid number 1 to 45, and domain of thiolase-like superfamily consisting of amino acid number 46 to 169. Thiolase superfamily is an important class of enzymes that belong to CoA dependent enzyme group. Thiolase has a key role in drug discovery so it's an interesting approach that *HMGS* has a thiolase like domain because HMG-CoA synthase itself use in cholesterol lowering drugs [12].

Features and visualization of proteoforms and interactive integration of HMG-CoA synthase sequence of *Brassica rapa* was analyzed through protter analysis. The analysis has revealed that query protein is found intracellular. No signal peptide has been predicted in the query protein and one residue (i.e. asparagine) was found belonging to N-glyco motif at position 140. N-linked glycosylation is the attachment of the sugar molecule oligosaccharide to the amide nitrogen of an asparagine (Asn) residue. This type of linkage is important for both the structure and function of some eukaryotic proteins.

Secondary structure prediction

The important concepts in secondary structure prediction are

S. no	Properties	Value
1	Number of amino acids	172
2	Molecular weight	19353.76
3	Theoretical PI	5.18
4	Ext. coefficient	29130
5	Instability index	36.40
6	Aliphatic index 76.05	76.05
7	Grand average of hydropathicity (GRAVY)	-0.309

Table 1: Physicochemical properties of HMG CoA synthase like gene predicted by ProtParam.

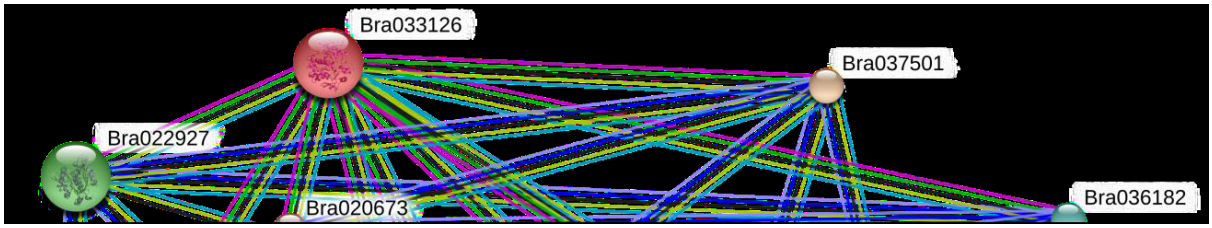

Figure 4: Networks are showing interactions of query protein sequence with different proteins. Different colours of nodes and strings present the protein-protein interactions.

identified as: residue conformational propensities, sequence edge effects, moments of hydrophobicity, position of insertions and deletions in aligned homologous sequence, moments of conservation, autocorrelation, residue ratios, secondary structure feedback effects and filtering. PSIPRED is a simple and accurate secondary structure prediction method, incorporating two feed-forward neural networks which perform an analysis on output obtained from PSI-BLAST (Position Specific Iterated - BLAST). The secondary structure of HMG-CoA synthase like protein predicted through PSIPRED. Six β sheets and three α helices were found in the predicted structure. 3D structure of *HMGS* describe by chang and his co-workers in 2015 that there were five layered core structure in following arrangement of α-β-α-β-α.

Homology modeling

Homology modelling, also known as comparative modelling of protein refers to constructing an atomic-resolution model of the "*target*" protein from its amino acid sequence and an experimental three-dimensional structure of a related homologous protein (the "*template*"). Modeller v9.16 was used for homology modelling and templates were selected on CPH server. Figure 5 was showing predicted 3D structure of HMG-CoA synthase protein from *Brassica rapa*.

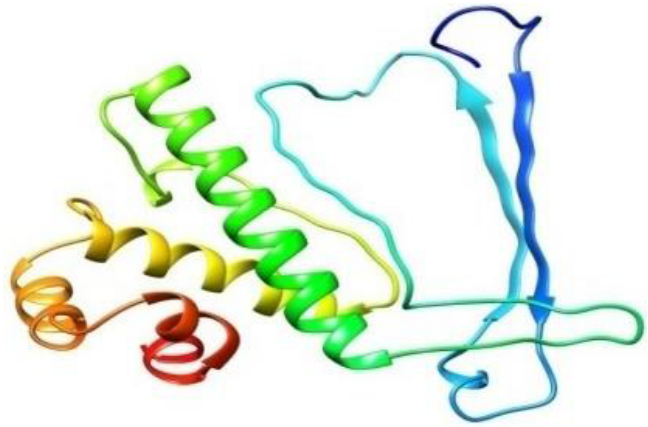

Figure 5: 3-D structure of HMG-CoA synthase from *Brassica rap*. β-strands, α-helix and coils are shown.

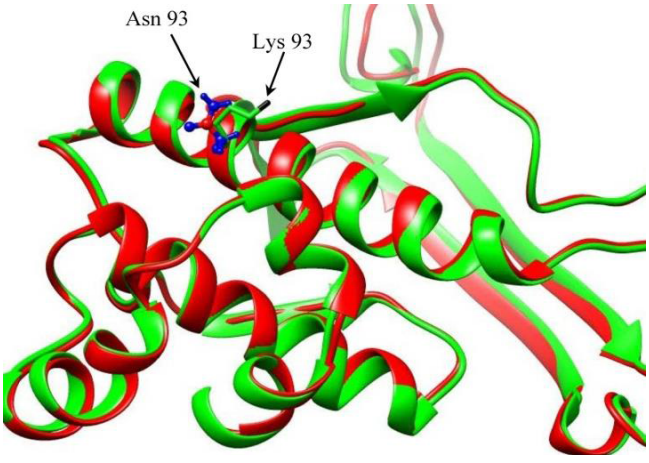

Figure 6: Superimpositions of predicted model. HMG-CoA from *B. rapa* (red) with HMG-CoA from *B. juncea* (Forest green).

Helix link helix coils were showed in 3D analysis of *HMGS* protein from *B. rapa*. The basic helix-loop-helix proteins have dimeric transcriptional factors that are found in almost all eukaryotes. Members of the bHLH superfamily have two highly conserved and functionally distinct domains, which together make a region of approximately 60 amino-acid residues. Furthermore, hydrophilic amino acids were present on the surface of the protein structure and hydrophobic amino acids residues were buried inside the structure [15]. When the 3D structure of HMG-CoA synthase from *B. rapa* was superimposed with its template (i.e., HMG-CoA synthase from *B. juncea*), RMSD of 0.652 Å was observed between both structures (Figure 6). In *B. juncea*, HMG-CoA synthase the lysine residue at position 93 was replaced with asparagine in *B. rapa* (shown in the form of atoms). It was revealed that there was 99.42% identity between 3D structures of HMG-CoA synthase proteins of *B. juncea* and *B. rapa.*

As Mustafa and his coworkers perform different bioinformatics tools, servers and softwares onGDF11 protein. Multiple tools for secondary structure prediction were used and the information about rabbit GDF11 given by these tools was compared. They were used PSIRPED and Modellerv 9.16 for 3D structure of GDF11 protein and gave homology modelling structure based on superimposition method.

Conclusion

In this study we have successfully annotated the mature segment of HMG-CoA synthase like gene and revealed information about protein-protein interaction. The study has also explored physicochemical nature of *HMGS*, secondary structure prediction through multiple tools and comparison of the information given by these tools. A phylogenetic analysis of nucleotide and protein was also conducted which revealed that our gene sequence have evolved from *Brassicacea* family with thiolase like superfamily domain.

Acknowledgements

I specially acknowledge to Dr. Ghulam Mustafa who guided me for the accomplishment of my work and highly acknowledge to Molecular Biochemistry Lab, Department of Biochemistry, University of Agriculture Faisalabad where all the wet work of this project was completed.

References

1. Apweiler R, Attwood TK, Bairoch A, Bateman A, Birney E, et al. (2001) The InterPro database, an integrated documentation resource for protein families, domains and functional sites. Nucleic Acids Res 29: 37-40.
2. Batool M, Khalid MH, Hassan MN, Yusuf HF (2011) Homology modeling of an antifungal metabolite plipastatin synthase from the *Bacillus subtilis* 168. J Bioinform 7: 384-387.
3. Besser RN, Eme UE, Ogbogu S (2009) Biodiversity and conservation of medicinal and Aromatic plants in Africa. Biotech Mol Biol Revi 3: 127-134.
4. Boukaftane T, Adeosun GO, Oduola TA, Avwioro GO, Oyeniyi MA (2005) Mechanism of action of *Jatropha gossypifolia* stem latex as a haemostatic agent. Eur J Gen Med 2: 140-143.
5. Buchan DW, Minneci F, Nugent TC, Bryson K, Jones DT (2013) Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic Acids Res 41: 349-357.
6. Caputi L, Aprea E (2011) Use of terpenoids as natural flavouring compounds in food industry. Rec Pat Food Nutri Agri 3: 9-16.
7. Chang J, Ning Y, Xu F, Cheng S, Li X (2015) Research advance of 3-hydroxy-3-methylglutaryl-coenzyme a synthase in plant isoprenoid biosynthesis. J Ani Plant Sci 25: 1441-1450.
8. Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12: 13-15.
9. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, et al. (2016) The Pfam

protein families database: Towards a more sustainable future. Nucleic Acids Res 4: 279- 285.

10. Fiser A, Sali A (2003) Modeller: generation and refinement of homology-based protein structure models. Methods In Enzymol 374: 461-491.
11. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, et al. (2005) Protein identification and analysis tools on the ExPASy server. John M Walker (ed) In: The Proteomics Protocols Handbook, Humana Press pp: 571-607.
12. Harijan RK (2017) Coenzyme-A dependent catalysis: An overview of thiolase superfamily enzymes and drug discovery. Res Chronic Dis 1: 17-19.
13. Holdgate GA, Ward WHJ, Taggart FM (2003) Molecular mechanism for inhibition of 3- hydroxy-3-methylglutaryl CoA (HMG-CoA) reductase by rosuvastatin. Biochem Soc Trans 31: 528-531.
14. Irshad F, Mushtaq Z, Akhter S (2017) Sequence analysis and comparative bioinformatics study of camelysin gene (*calY*) isolated from *Bacillus thuringiensis.* Biochem Gene.
15. Jones S (2004) An overview of the basic helix-loop-helix proteins. Genome Biol 5: 1-6.
16. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M (2012) Geneious basics: An integrated and extendable desktop software platform for the organization and analysis of sequence data. J Bioinform 28: 1647–1649.
17. Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16: 111-120.
18. Liao P, Wang H, Wang M, Hsiao AS, Bach TJ, et al. (2014) Transgenic tobacco overexpressing *Brassica juncea* HMG-CoA synthase 1 shows increased plant growth, Pod Size and Seed Yield. J Biol Chem 9: 1-15.
19. Ma B, Elkayam T, Wolfson H, Nussinov R (2003) Protein–protein interactions: Structurally conserved residues distinguish between binding sites and exposed protein surfaces. Proc Natl Acad Sci 100: 5772-5777.
20. Mustafa G, Iqbal J, Hassan M, Jamil A (2017) Bioinformatics characterization of growth differentiation factor 11 of *Oryctolagus cuniculus.* J Chem Soc Pak 39: 1089-1095.
21. Nielsen M, Lundegaard C, Lund O, Petersen TN (2010) CPHmodels-3.0-remote homology modeling using structure-guided sequence profiles. Nucleic Acids Res 38: 576-581.
22. Olivier LM, Kovacs W, Masuda K, Keller GA, Krisans SK (2000) Identification of peroxisomal targeting signals in cholesterol biosynthetic enzymes: AA-CoA thiolase, HMG-CoA synthase, MPPD and FPP synthase. J Lipid Res 41: 1921-1935.
23. Omasits U, Ahrens CH, Müller S, Wollscheid B (2013) Protter: interactive protein feature visualization and integration with experimental proteomic data. J Bioinform 30: 884-886.
24. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, et al. (2004) UCSF Chimera-a visualization system for exploratory research and analysis. J Comput Chem 25: 1605-1612.
25. Pojer F, Ferre JL, Richard SB, Nagegowda DA, Chye ML, et al. (2006) Structural basis for the design of potent and species-specific inhibitors of 3-hydroxy-3-methylglutaryl CoA synthases. Proc Natl Acad Sci 103: 11491–11496.
26. Saitou N, Nei M (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. J Mol Biol Evol 4: 406-425.
27. Schwarz R, Dayhoff M (1979) Matrices for detecting distant relationships. In Dayhoff M (ed), Atlas of protein sequences. National Biom Res Foun 7: 353-358.
28. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, et al. (2017) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res 45: 362-368.
29. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance and maximum parsimony methods. Mol Biol Evol 28: 2731-2739.
30. Theisen MJ, Misra I, Saadat D, Campobasso N, Miziorko HM, et al. (2004) 3-hydroxy-3-methylglutaryl-CoA synthase intermediate complex observed in ‘‘real-time. Proc Natl Acad Sci 101: 6442–16447.