Plant Omics Journal POJ

POJ 8(1):1-8 (2015) ISSN:1836-3644

Comparative computational analysis of stigmasterol biosynthetic genes and proteins in plants

Bawankar Raksha, Rajendrarao Priya, Ramamoorthy Siva and Subramanian Babu

School of Bio Sciences and Technology, VIT University, Vellore 632014, India

Corresponding author: babu.s@vit.ac.in

Abstract

Plant stigmasterol synthesis involves the partaking of many genes. In the present study, we performed computational comparison of the genes *smt2*, *smo2*, *ste1*, *dwf5*, *dwf1* and c-22 sterol desaturase involved in the phytosterol biosynthetic pathway in different plants. Eleven plants belonging to different botanical families were selected and the gene sequence was retrieved from the KEGG pathway database. Nucleotide sequence homology was observed by Clustal W analysis and phylogeny was studied using MEGA 5 software. Clustal W and MEGA 5 analysis was done for amino acid sequence also. The maximum conserved regions were used in Prosite scan analysis. To determine the domains in protein sequence of the aforesaid genes in eleven plants, Pfam domain database search was performed. Phyre 2 software was used to develop 3D protein models. All the monocots used in the study belong to *Poaceae* family and hence in our phylogenetic analysis of phytosterol biosynthetic genes showed 90% sequence homology among the plants with same point of origin. The dicot plants chosen for the study belong to different families and hence the genes showed a homology percentile of less than 80.Pfam studies revealed SMT2 protein with a Methyl transf 11 and Sterol MTC domains. SMO2 and STE-1 proteins showed a common FA hydroxylase domain. DWF-5 was found to have domain structure of ERG4/ERG24 family, whereas, DWF-1 showed presence of FAD binding 4 domain. P450 structural domain was represented by sterol c-22 desaturase proteins. The 3D structure prediction revealed structural similarity among all eleven plants for protein involved in the plant stigmasterol biosynthesis.

Keywords: Biosynthetic genes; Phylogenetic analysis; Plants; Proteins; Stigmasterol.
Abbreviations: *smt2*_24-methylene sterol C- methyl transferase, *smo2*_4-alpha methyl- delta-7-sterol-4 alpha methyl sterol oxidase, *ste*-1_lathosterol oxidase, *dwf5*_7-dehydrocholesterol reductase, *dwf1*_delta 24- sterol reductase, *sterol c-22 desaturase*_ cytochrome P450 family 710, subfamily A, FA hydroxylase_Fatty acid hydroxylase, FAD_Fatty acid domain.

Introduction

In higher plants the major lipid constituent of the plasma membrane are sterols (Haines, 2001) and its minor proportion serves as precursor to steroid derivatives (Clouse, 2001). The presence of phytosterols has been observed mostly in food products and has structural similarity to cholesterol (Normen et al., 2000). Amongst all phytosterols, stigmasterol and sitosterol are major components of sterol profiles of plant species (Schaller, 2004; Schrick, 2002). In this study, we compared stigmasterol biosynthetic genes and proteins of eleven different plants. The plant stigmasterol biosynthesis pathway initiates from farnesyl diphosphate farnesyl transferase. This microsomal enzyme catalyses the first committed step in the biosynthesis of sterols. The synthesis of plant stigmasterol initiates from cycloartenol and its transformation to delta-5 sterols in higher plants is a multistep process (Benveniste, 1986), followed by sterol 24-C-methyl transferase in the presence of glutathione which is a necessitate component and acts on 24 (25)- double bond in the side chain of sterol (Moore and Gaylor, 1969; Venkatramesh et al., 1996, Tong et al., 1997). After 24-methylene cycloartenol, phytosterol biosynthesis pathway is linear, until it reaches to 24-methylene lophenol, since it gets divided into two main pathways that escort to 24- methyl and 24-ethyl sterols (Taton and Rahier, 1991). The 24-methylenesterol C-methyl transferase (SMT2) is the second methylation step of plant sterol biosynthesis which considers 24(28)-methylene lophenol as the substrate (Rahier et al., 1986, Bouvier-Nave et al., 1997). It has the capability to synthesize 24-ethyl sterols from a 24-methylene sterol precursor (Benveniste, 1986; Bouvier-Nave et al., 1997). It has been already proved that CYP710A catalyses the C-22 desaturase reaction for *in vitro* production of stigmasterol from β-sitosterol (Morikawa et al., 2006 a). The 24-ethyledenelophenol produces delta-7-avenasterol with the help of enzyme SMO2 (4-alpha-methyl-delta 7-sterol-4 alpha-methyl oxidase). SMT2 and SMO2 enzymes assist in the establishment for the ratio of campesterol to sitosterol it helps in the plant development and membrane integrity (Schaller, 1998; Schaeffer et al., 2001; Carland et al., 2002; Hase et al., 2005; Schaller, 2004). The STE1 enzyme (lathosterol oxidase) catalyzes the introduction of a C-5 double bond in to the B ring of delta- 7-sterols to yield the corresponding delta-5, 7-sterols (Nishino et al., 1997). Isofucosterol synthesis occurs when 5-hydroavenosterol acts as a substrate for the enzyme DWF5 (Choe et al., 2000). The synthesis of end sterol *i.e.* the sitosterol is depend on the enzyme DWF1, when combines with synthesized isofucosterol. DWF1 acts as a biosynthetic enzyme, catalyzing C-24 reduction and it is a key step for plant sterol biosynthesis (Klahre et al., 1998). In the plant kingdom, there

is a broad distribution of β-sitosterol. P450, CYP710A family is a sturdy entrant of the C-22 desaturase among widely diversified plant cytochromes (Nelson et al., 2004). The CYP710A displays C-22 desaturase activity with β-sitosterol to produce stigmasterol (Morikawa et al., 2006 b). Stigmasterol is a therapeutic product which is used as a precursor for the synthesis of hormone (Sundaram and Djerassi, 1997), vit.D_3 (Kametani and Furuyama, 1987) and inhibit the absorption of cholesterol (Gabay et al., 2010). Stigmasterol exhibits elevated therapeutic properties and can be used as a healing agent for many disorders. The present study includes eleven different plants and six different genes *i.e. smt2* (24-methylene sterol C- methyl transferase), *smo2* (4-alpha methyl- delta7-sterol-4 alpha methyl sterol oxidase), *ste1* (lathosterol oxidase), *dwf5* (7-dehydrocholesterol reductase), *dwf1* (delta 24- sterol reductase) and sterol c-22 desaturase (cytochrome P450 family 710, subfamily A) involved in stigmasterol biosynthesis. The objective of this computational analysis is to study the similarities and differences between stigmasterol biosynthetic genes of crops and plants of different botanical families from the functional and evolutionary perspective. To our knowledge, such comparative analysis of stigmasterol biosynthetic genes among plant species has not been reported previously.

Results and Discussion

Plants and stigmasterol biosynthetic pathway

The details of the plants included in the study are given in Table 1. The stigmasterol biosynthetic pathway in plants is shown in Fig. 1.

Multiple sequence alignment and phylogenetic evolution studies

The results of multiple sequence alignment are shown in Supplementary Table 1. The results of phylogenetic analysis are shown in Fig. 2. Among monocots we observed a high percentage of homology (90%) in gene sequences since they belong to *Poaceae* family and the phylogenetic evolution also reveals their similarity in origin. But in the case of dicots there is varied origin of evolution for each plant, as they belong to varied family and the homology of gene sequences was less than 80%. The *smt-2* gene of *Arabidopsis* is known to be involved in abiotic stress tolerance leading to leakage of cellular nutrients and ions into apoplast (Senthil Kumar et al., 2013). The gene sequence showed a highest homology with *R. communis* (76.15%). For all other genes the sequence homology was less comparing *Arabidopsis* with other plants. The *smo2*, *dwf-5* and *dwf-1* gene of *C. sativus* showed more than 75% sequence homology with *R.communis*. The first five gene of the pathway (*smt-2*, *smo2*, *ste1*, *dwf5* and *dwf1*) showed 82.04, 86.17, 79.78, 82.80 and 84.93% sequence homology between *R. communis* and *P. trichocarpa*. The smo2 gene sequence of *S. lycopersicum* has shown 78.52% homology with *C. arietinum*. The *dwf1* gene sequence of *R. communis* has 80.08%, 79.28% and 81.78% homology with corresponding *dwf-1* gene of *P. trichocarpa*, *R. communis* and *V. vinifera*, respectively. Among monocot plants the gene sequences has shown more than 90% homology since all of plant belonged to *Poaceae* family. The percentage was only less than 75% homology in gene sequences between monocots and dicots. Arabidopsis plants are reported to contain 6% stigmasterol (Benveniste et al., 2004). It has been reported that phytosterol component of castor seeds have antifertility and contraceptive efficacy properties (Nithya et al., 2011). Han et al. (2008) reported 1.9 mg of stigmasterol per part of tomato. Whitaker and Gapper (2008) reported the increase in stigmasterol during ripening stage corresponding to the increase in the expression of *sterol c-22 desaturase* gene. Adeyeye and Adesina (2013) reported 5.75-8.46 mg stigmasterol per 100 g of edible part of *S. bicolor*. Kislichenko et al. (2009) quantified 0.398 mg/kg of foxtail millet (*S. italica*). Among the genes of stigmasterol biosynthesis, sterol desaturase was found to be less similar among the plants in terms of sequence homology indicating its diversification towards varied amount of stigmasterol in the plants. The observation of Whitaker and Gapper (2008) of a correlation between stigmasterol content and sterol desaturase gene expression in tomato is supportive to designate this gene as a key gene for stigmasterol accumulation in plants. The phylogenetic analysis results of stigmasterol biosynthetic genes are shown in Fig.2. For all genes, the tree show independent evolution of monocots compared to dicots. Fig. 2. (a) represents the evolution of *smt2* gene. In this phylogenetic tree, the sum of branch length was 1.37785575. In the bootstrap test the percentage of replicating trees for eleven plant sequences were represented next to the branches. The *B. dystachyon*, *S. bicolor*, *O. sativa* and *S. italica* are showing more homology for *smt2* gene. Among dicot species, *V. vinifera* is getting separately branched away from all remaining six plants. Fig. 2. (b) shows phylogenesis of *smo2* gene. The sum of branch length was 1.29235020 and the unit scale was 0.05. Among monocots *O. sativa* is showing high divergent evolution compared to others. Among dicots, *A. thaliana* and *C. sativus* are branching individually away from other five plants. The *smo2* gene of *C. arietinum* can be observed as individual branch before *R.* communis and *P. trichocarpa*. Fig. 2. (c) represents the phylogenetic tree for *ste1*. This tree has shown the sum of branch length as 1.64809938 and unit scale of 0.05 was used. All the monocot plants are showing a high percentage for evolutionary similarity for this gene. On the other hand among dicots, *V. vinifera* is showing separate diverge branch of evolution. *C. sativus* is showing separate branch of evolution before *S. lycopersicum* and *C. arietinum*. Fig. 2. (d) represents phylogenetic tree for *dwf5* gene. The sum of branch length was 12.29730565. Among monocots, comparing to *S. italica*, *O. sativa*, *S. bicolor* and *B. distachyon* are sharing a similar percentage of evolution pattern for *dwf5* gene. Among dicots, *R. communis*, *V. vinifera*, *S. lycopersicum* and *P. trichocarpa* are showing more homology and found to have evolved from third branch. *C. sativus* and *C. arietinum* are present in second branch among dicots. *A. thaliana* has evolved as individual diverge branch representing its uniqueness in evolution of *dwf5* gene. Fig. 2. (e) represents a phylogenetic tree for *dwf1* gene. The sum of branch length was 1.26096286. *S. italica* and *B. distachyon* have shown homology as compared to other two monocots. Among dicot plants, *A. thaliana* and *S. lycopersicum* are sharing same ancestor branch and showing more homology. *C. arietinum* has shown independent evolutionary pattern for this gene. Fig. 2. (f) represents the phylogenetic tree of *sterol* c-22 *desaturase* gene. The sum of branch length obtained as 1.87907341. *S. bicolor* and *S. italica* have evolved from the same ancestor branch and showing homology. *O. sativa* has evolved from the second branch followed by *B. distachyon* which has evolved individually from the third branch. Among the dicots, *A. thaliana* has shown individual branch of sterol desaturase gene evolution which has originated from the first branch.

Table 1. Preface about the plants under in the study.

S.No.	Scientific name	Common Name	Family	Monocot/Dicot
1.	*Solanum lycopersicum*	Tomato	*Solanaceae*	Dicot
2.	*Arabidopsis thaliana*	Thale cress	*Brassicaceae*	Dicot
3.	*Populus trichocarpa*	Western balsam poplar	*Salicaceae*	Dicot
4.	*Sorghum bicolor*	*Sorghum*; broomcorn; milo	*Poaceae*	Monocot
5.	*Cucumis sativus*	Cucumber	*Cucurbitaceae*	Dicot
6.	*Ricinus communis*	Castor	*Euphorbiaceae*	Dicot
7.	*Vitis vinifera*	Grape wine	*Vitaceae*	Dicot
8.	*Cicer arietinum*	Chick pea	*Fabaceae*	Dicot
9.	*Setaria italic*	Foxtail millet	*Poaceae*	Monocot
10.	*Oryza sativa*	Asian Rice	*Poaceae*	Monocot
11.	*Brachypodium distachyon*	Purple false brome	*Poaceae*	Monocot

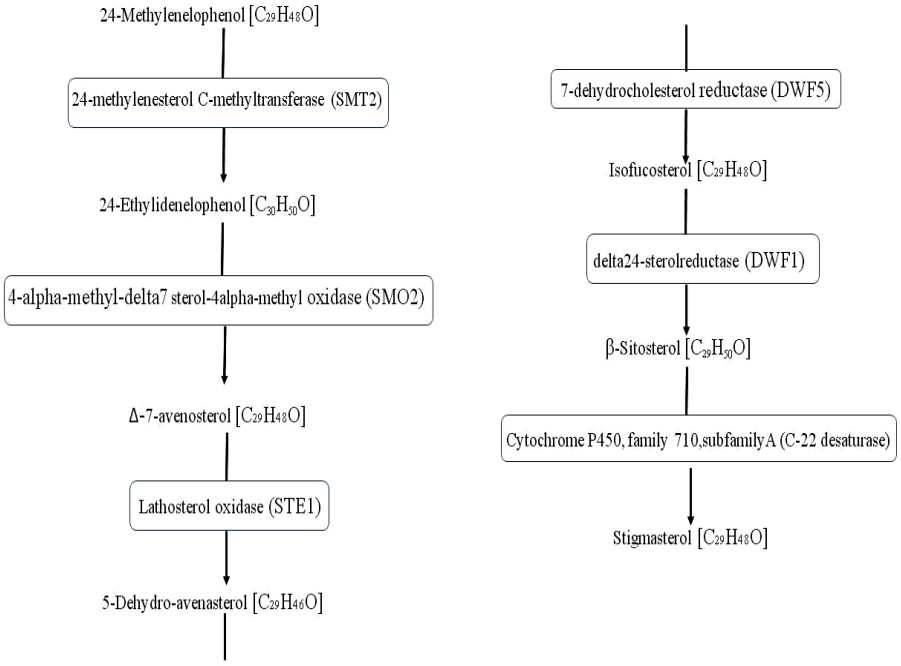

Fig. 1. Stigmasterol biosynthetic pathway in plants showing the involvement of enzymes and intermediate metabolites

The evolution of *V. vinifera* and *C. arietinum* are genes have shown independent second and third evolutionary branch respectively. *R. communis* and *P. trichocarpa* as well as *C. sativus* and *S. lycopersicum* were found to share a common ancestor.

Pfam domain annotation of six stigmasterol biosynthetic proteins and their structural analysis

For the domain annotation using amino acid sequences of stigmasterol biosynthetic proteins, Pfam (http://pfam.sanger.ac.uk/) database was used to scan Pfam A and Pfam B domain models. The results are presented in supplementary Table 2. The SMT2 protein exhibited two domains *viz.* a. methyl transf 11 and b. sterol MTC. Methyl transferase 11 is known to be involved in the methyl transfer from the S--adenosyl-L-methionine (SAM) to either nitrogen, oxygen or carbon atoms and this is known in plants as well as in bacteria and in mammals. The SMT1 gene encodes *S*-adenosylmethionine–dependent C-24 SMT which leads to addition of single methyl group. SMT1 is capable to catalyze both alkylation steps which exhibit substrate specificity of enzymes in sterol biosynthesis (Diener et al., 2000). SMO2 and STE-1 proteins are exhibited thepresence of domain for fatty acid hydroxylase (FA hydroxylase). These are mainly presented in sphingolipids and they are synthesized by a shingolipid fatty acid 2- hydroxylase (FAH) (Nagano et al., 2012). Pata et al. (2010) indicated FA hydroxylase in plants like *O. sativa*, *Z. mays*, *G. max*, *S.* bicolor and *V. vinifera* in addition to *Arabidopsis*. DWF-5 includes the domain structure of ERG4/ERG24 family. Few fungal enzyme like C-24(28) sterol reductase and C-14 sterol reductase are said to be involved in ergosterol biosynthesis and they are present at the budding stage of yeast. Sterol biosynthesis pathways have common steps in fungi, plants, and animals; the only variation is that fungi have an additional branch that adds a methylidene/methyl group at position C-24 and desaturates position C-22 to form ergosterol (Bagnat et al., 2001; Umebayashi

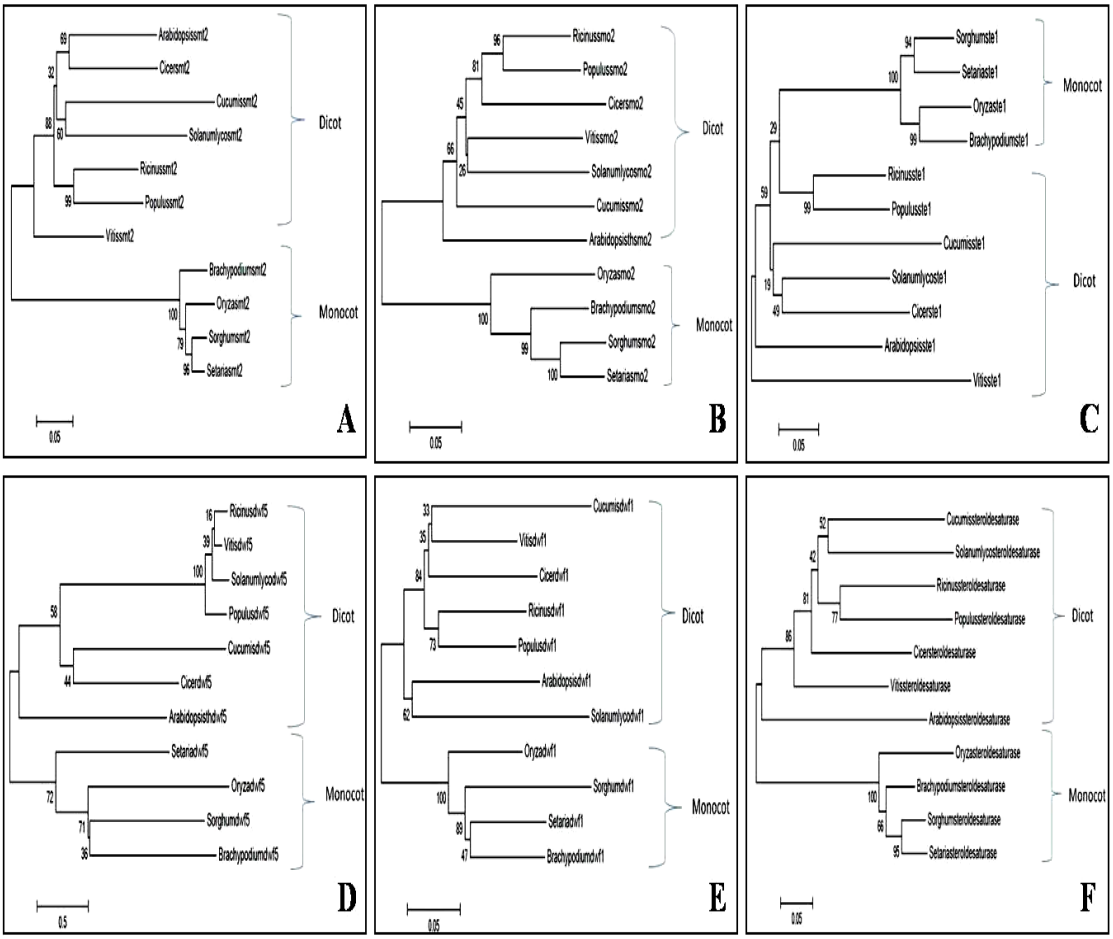

Fig 2. Phylogenetic trees showing the evolutionary relationships of stigmasterol biosynthetic genes among the different plants A) *smt2* B) *smo2 C*) *ste1* D) *dwf5* E) *dwf1* F) *sterol c-22 desaturase*.

and Nakano, 2003). DWF-1 represented the presence of FAD binding 4 domain. This flavin adenine dinucleotide consists of adenosine monophosphate (AMP) linked to flavin mononucleotide (FMN) by a pyrophosphate bond (Dym and Eisenberg, 2001). P450 structural domain was represented by sterol c-22 desaturase proteins. This enzyme catalyze and the monooxygenetion of group from the precursor. Over 3700 cytochrome P450 genes have been identified, among them few were identified in plants (Nelson et al., 2004).

Protein motif PROSITE SCAN of aligned amino acid sequences

Several studies which involve sequence comparison include the analysis of protein motif and the presence of orthologous groups. Protein motif was selected from the alignment result of amino acid sequences with the Clustal W from MEGA 5.2.2 and further it was subjected to PROSITE SCAN (http://prosite.expasy.org/scanprosite/). The results obtained in the form of patterns and the hits at the motif region are presented in supplementary Table 3. and Fig. 3.Prosite is a database of biologically significant protein site and patterns which helps to identify the known family of proteins for the studied sequence with the help of computational tools. The specific characters with some biological properties which are shared by the protein families denotes the term "PATTERN" for that sequence (Comet and Henry, 2002). In the interpretation of PROSITE database, several other tools are involved, for example Profile Scan uses the method of Gribskov et al. (1988) for the analysis of structural and sequence motifs in a protein sequence. This also helps to scan sequences from Swiss Prot or TrEMBL (Bairoch and Apweiler, 1997). In the representation of Prosite scan motif, upper case represents match positions, lower case is the insert positions, and the '-' symbol represents deletions relative to the matching profile. There is a distribution of data as false positive, true positive, false positive, false negative, potential sequence and unknown sequence. The data output in our study have shown for true positive and false negatives. True positives are the sequences which represent the regular expression modeling pattern which belong to biological family associated with available pattern. In *vice versa* the false negatives are those which do not represent the regular expression. Class I SAM-dependent methyltransferases family profiles depicts that methyltransferases are an important class of enzymes which are present in every life form. They are responsible for the transfer of a methyl group mostly from S-adenosyl L-methionine (SAM or AdoMet) to a nucleophilic acceptor such as nitrogen, oxygen, sulfur or carbon leading to S-adenosyl-L-homocysteine (AdoHcy) and a methylated molecule. Methyltransferases are involved in many necessary cellular processes which includes biosynthesis, signal transduction, protein repair, chromatin

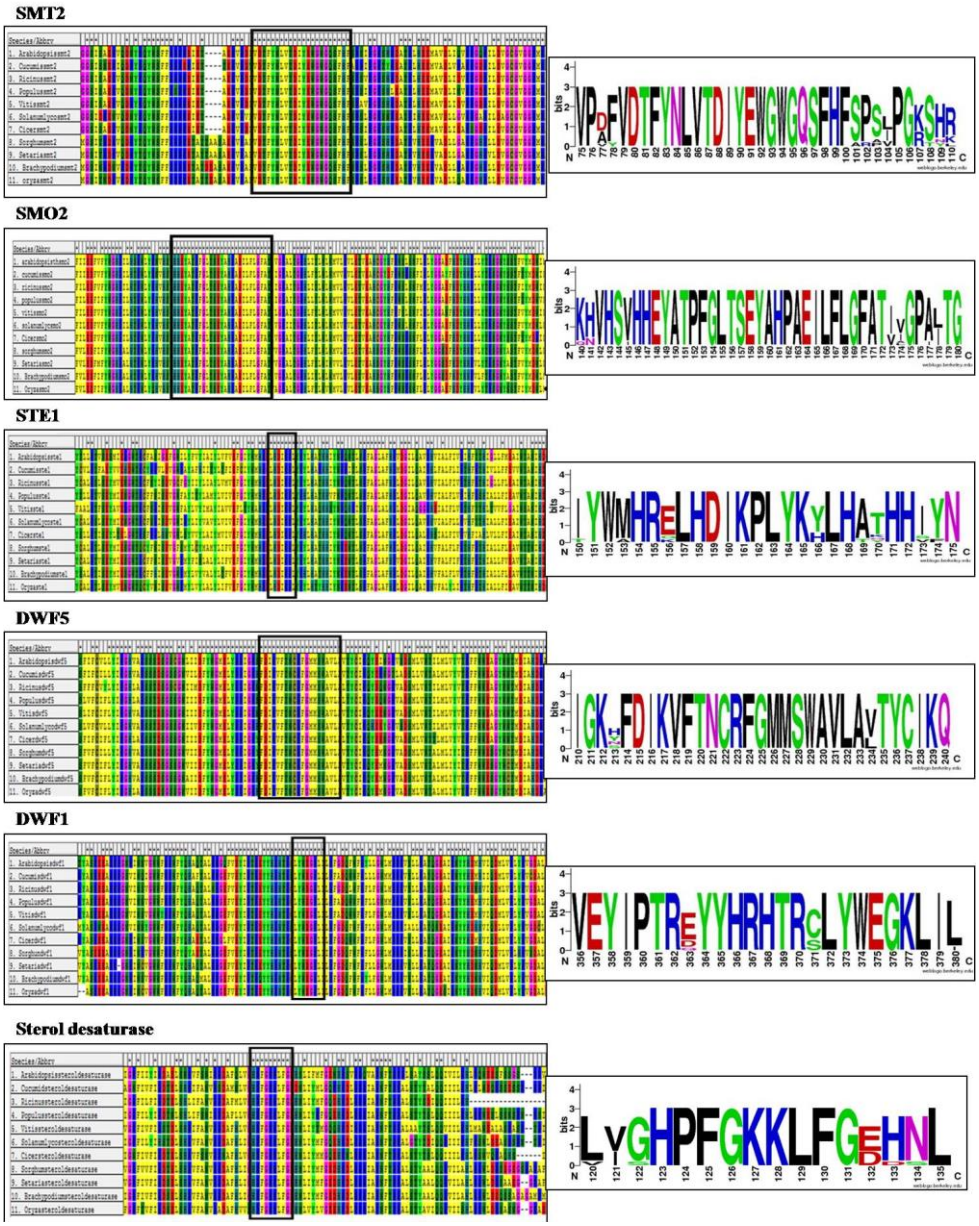

Fig 3. ClustalW analysis amino acid sequences of stigmasterol biosynthetic proteins and selection of motif.

regulation and gene silencing (Kozbial and Mushegian, 2005; Schubert et al., 2003; Wlodarski et al., 2011). SAM-dependent methyltransferase Erg6/SMT-type domain profile shows that sequences in UniProtKB/Swiss-Prot are known to belong to this class: 26 which are detected by PS51685:26(true positives) and undetected by PS51685: 0 (false negative or 'partial'). SMO2 and STE-1, have shown no hit in the PROSITE motifs, which shows the lack of database entry for this specific protein. DWF5 belong to a conserved family of sterol reductases which act by reducing double bonds (Lai et al., 1994; Holmer et al., 1998), STEROL_REDUCT_1, PS01017; Sterol reductase family signature 1 (PATTERN) gives Consensus pattern of G-x(2)-[LIVMF]-[IFYH]-[DN]-x-[FYWM]-x-G-x(2)-[LF]-[NY]-P-[RQ]. Sequences in UniProtKB/Swiss-Prot are known to belong to this class: 27 which are detected by PS01017: 25 (true positives) undetected by PS01017: 2 (2 false negatives and 0 'partial').

And in the case of PS01018 Consensus pattern is [LIVM](2)-[LIVMFT]-[HWD]-R-x(2)-R-D-x(3)-C-x(2)-K-Y-[GK]-x(2)-[FW]-x(2)-Y. Sequences in UniProtKB/Swiss-Prot are known to belong to this class: 27 which are detected by PS01018: 25 (true positives) and undetected by PS01018: 2 (2 false negatives and 0 'partial'). When we analyzed the Scan prosite motif hit for DWF1, PCMH-type FAD-binding domain profile reveals that flavoenzymes are having the capacity to catalyze biochemical reactions and they are involved in the dehydrogenation of a variety of metabolites, in electron transfer from and to redox centres, in light emission, in the activation of oxygen for oxidation and hydroxylation reactions (Mathews 1991; Fraaije and Mattevi, 2000). Sequences in UniProtKB/Swiss-Prot known to belong to this class: 628: which are detected by PS51387:628(true positives) and undetected by PS51387: 0 (false negative or 'partial'). CYP710A which encodes sterol desaturase protein, showed two hits for PROSITE scan motif hit. Cytochrome

P450's (Nebert and Gonzalez, 1987; Coon et al.,1992) are a group of enzymes which are involved in the oxidative metabolism natural compounds (such as steroids, fatty acids, prostaglandins and leukotrienes) as well as drugs, carcinogens and mutagens. Cytochrome P450 cysteine heme-iron ligand signature (PATTERN) depicts the consensus pattern of [FW]-[SGNH]-x-[GD]-{F}-[RKHPT]-{P}-C-[LIVMFAP]-[GAD]. In this C is the heme iron ligand and the sequences in UniProtKB/Swiss-Prot are known to belong to this class:1035 which are detected by PS00086:955 (true positives) and undetected by PS00086: 80 (70 false negatives and 10'partials')

3D Protein model construction using Phyre2

With the reference to known protein structures the normal mode modeling by Phyre2 has produced a set of potential 3D models of proteins of our interest. The results are presented in supplementary Fig. 1. Rasmol is the mode of structure viewer used which has given the 3D pattern of protein. The structural prediction shows that all eleven plants for each protein are visually similar. We have used all conserved regions from sequence alignment for protein of all plants and it has shown similar structural appearance. This approach has the capacity to produce model with about 70% of the domains in a typical genome, and this is possible because of the use of remote homology detection technique such as profile-profile and HMM-HMM matching. In addition, it is also possible to confidently and correctly detect the homology even if it has sequence identity as low as 15% between a pair of proteins (Kelley and Sternberg, 2009).

Materials and methods

Source of data

Sterol biosynthetic pathway was retrieved from KEGG (Kyoto Encyclopedia of Genes and Genomes; pathway, (http://www.kegg.jp/kegg-bin/search_pathway_) (Kanehisa, 2014). The nucleotide and amino acid sequences of *smt2* (24-methylene sterol C- methyl transferase), *smo2* (4-alpha methyl- delta-7-sterol-4 alpha methyl sterol oxidase),ste-1 (lathosterol oxidase), dwf5 (7-dehydrocholesterol reductase), *dwf1* (delta 24- sterol reductase), *sterol c-22 desaturase* (cytochrome P450 family 710, subfamily A) for eleven different plants, *Solanum lycopersicum* (*Solanaceae*), *Arabidopsis thaliana* (*Brassicaceae*), *Populus trichocarpa* (*Salicaceae*), *Sorghum bicolor* (*Poaceae*), *Cucumis sativus* (*Cucurbitaceae*), *Ricinus communis* (*Euphorbiaceae*), *Vitis vinifera* (*Vitaceae*), *Cicer arietinum* (*Fabaceae*), *Setaria italica* (*Poaceae*), *Oryza sativa* (*Poaceae*), *Brachypodium distachyon* (*Poaceae*) were obtained from this database.

ClustalW and MEGA 5.2.2 analysis

Multiple sequence alignment was performed by using ClustalW (http://www.genome.jp/tools/clustalw/). ClustalW has given the homology score, which helped to analyze the percentage of similarity among the plants. *smt2*, *smo2*, *ste1*, *dwf5*, *dwf1* and *sterol c-22 desaturase* genes based on sequence alignments, phylogenetic trees were prepared using MEGA5.2.2 (http://www.megasoftware.net/) with the bootstrap value 1000 in neighbor - joining method (Tamura et al., 2011).

Pfam domain annotation and protein motif scan

We analyzed the protein sequences for Pfam matches (http://pfam.sanger.ac.uk), which provides data for identification of domains that occur within proteins and thus highlighting their function (Punta et al., 2012). The most conserved region *i.e.* protein motif was selected from the aligned amino acid sequences of the proteins encoded by the stigmasterol biosynthetic genes among eleven different plants and it was represented with the help of Weblogo (Crooks et al., 2004; Schneider and Stephens, 1990) for the detection of the PROSITE signature matches in protein sequences, SCAN PROSITE (http://prosite.expasy.org/scanprosite/) (Castro et al., 2006).

3D Structure prediction of proteins

Protein homology/analogy recognition engine, (http://www.sbg.bio.ic.ac.uk/phyre2/html/page) (Kelley and Sternberg, 2009), was used for the protein structure prediction which includes the template based homology modeling or fold-recognition (Qian et al., 2007). We have submitted amino acid sequences retrieved from KEGG pathway database and a non-redundant fold library was constructed using these sequences. Sequences in the fold library were scanned against a non-redundant sequence database and a Hidden Markov Model (HMM) for each known structure. This fold library also consists of known and predicted secondary structures for all stored protein sequences. HMM was created by scanning the submitted protein sequences against the non-redundant sequence database. The close and remote sequence homologues were collected using PSI Blast, and an alignment was constructed, following this secondary structure prediction was completed. The profile HMM and the secondary structure were then used to scan the fold library using HMMe HMM matching. This alignment process returned a score on which all alignments were ranked, and an E-value was generated. The twenty top scoring matches were then used to generate full 3-D models of each sequence as per the procedure of Kelley and Sternberg (2009).

Conclusion

Although some attempts were made to over-express the CYP710A gene in *Arabidopsis* and tomato to prove its function in relation to the overall understanding of the sitosterol-stigmasterol biosynthesis, our knowledge on diversity of the genes involved in biosynthesis is still incomplete. As a preliminary step towards complete characterization of diversified sterol biosynthesis in plants, we attempted to compare at the gene and protein level. The similarities and differences observed among stigmasterol biosynthetic genes of crops and plants of different botanical families have given important molecular clues to understand the evolution of this pathway. To our knowledge, such comparative analysis of stigmasterol biosynthetic genes among plant species has not been reported previously. The results of our computational analysis warrant further *in vitro* and *in planta* research to understand the functional diversity of the sterol biosynthesis in plant species.

Acknowledgement

The support and the facilities offered by the management of VIT University, Vellore, Tamilnadu to carry out this research is gratefully acknowledged.

References

Adeyeye EI, Adesina AJ (2013) Enhanchement of lipid quality of raw guinea corn (*Sorghum bicolor*) grains through germination and steeping. Open J Analyt Chem Res. 1(1):5-17.

Bagnat M, Chang A, Simons K (2001) Plasma membrane proton ATPase Pma1p requires raft association for surface delivery in yeast. Mol Biol Cell. 12:4129–38.

Bairoch A, Apweiler R (1997) The swiss-prot protein sequence data bank and its supplement trembl. Nucleic Acids Res. 25 (1):31–36.

Benveniste P (1986) Sterol biosynthesis. Annu Rev Plant Physiol. 37:275-307.

Benveniste P (2004) Biosynthesis and accumulation of sterols. Annu Rev Plant Biol. 55:429–457.

Bouvier-Nave P, Husselstein T, Desprez T, Benveniste P (1997) Identification of cDNAs encoding sterol methyl transferases involved in the second methylation step of plant sterol biosynthesis. Eur J Biochem. 246:518-29.

Carland FM, Fujioka S, Takatsuto S, Yoshida S, Nelson T (2002) The identification of CVP1 reveals a role for sterols in vascular patterning. Plant Cell. 14:2045-2058.

Castro DE, Sigrist CJA, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, Bairoch A, Hulo N (2006) ScanProsite: detection of PROSITE signature matches and ProRule associated functional and structural residues in proteins. Nucleic Acids Res. 1:34 (Web Server issue):W362-5.

Choe S, Tanaka A, Noguchi T, Fujioka S, Takatsuto S, Ross AS, Tax FE, Yoshida S, Feldmann KA (2000) Lesions in the sterol delta 7 reductase gene of Arabidopsis cause dwarfism due to a block in brassinosteroid biosynthesis. Plant J. 21(5):431-443.

Clouse SD (2001) Integration of light and brassinosteroid signals in etiolated seedling growth. Trends Plant Sci. 10:443-5.

Comet JP, Henry J (2002) Pairwise sequence alignment using a PROSITE pattern-derived similarity score. Comput Chem. 26:421–436.

Coon MJ, Ding XX, Pernecky SJ, Vaz AD (1992) Cytochrome P450: progress and predictions. FASEB J. 6:669-673.

Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: A sequence logo generator. Genome Res. 14:1188-1190.

Diener AC, Li H, Zhou Wen-xu, Whoriskey WJ, Nes WD, Fink GR (2000) Sterol MTC: sterol methyltransferase 1 controls the level of cholesterol in plants. Plant Cell. 12:853–870.

Dym O, Eisenberg D (2001) Sequence-structure analysis of FAD-containing proteins. Protein Sci. 10:1712–1728.

Fraaije MW, Mattevi A (2000) Flavoenzymes: diverse catalysts with recurrent features. Trends Biochem Sci. 25:126-132.

Gabay O, Sanchez C, Salvat C, Chevy F, Breton M, Nourissat G, Wolf C, Jacques C, Berenbaun F (2010) Stigmasterol: a phytosterol with potential anti-osteoarthritic properties. Osteoarthr Cartil. 18(1):106-116.

Gribskov M, Homyak M, Edenfield J, Eisenberg D (1988) Profile scanning for three-dimensional structural patterns in protein sequences. Comput Appl Biosci. 4 (1):61–66.

Haines TH (2001) Do sterols reduce proton and sodium leaks through lipid bilayers?. Prog Lipid Res. 40:299-324.

Han JH, Yang YX, Feng MY (2008) Contents of phytosterols in vegetables and fruits commonly consumed in china. Biomed Environ Sci. 21(6): 449-453.

Hase Y, Fujioka S, Yoshida S, Sun S, Umeda M, Tanaka A (2005) Ectopic endore duplication caused by sterol alternation results in serrated petals in Arabidopsis. J Exp Bot. 56: 1263-1268.

Holmer L, Pezhman A, Worman HJ (1998) The human lamin B receptor/sterol reductase multigene family. Genomics. 54:469-476.

Kametani T, Furuyama H (1987) Synthesis of vitamin D_3 and related compounds. Med Res Rev. 7(2):147-171.

Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi1M, Tanabe M (2014) Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. 42: Database issue D199–D205 doi:10.1093/nar/gkt1076.

Kelley LA, Sternberg MJE (2009) Protein structure prediction on the web: a case study using the Phyre server. Nat Protocol. 4:3.

Kislichenko VS, Omel′chenko ZI, Vel′ma VV (2009) Sterols from *Setaria italica* grain. Chem Nat Compound. 45(2).

Klahre U, Noguchi T, Fujioka S, Takatsuto S, Yokota T, Nomura T, Yoshida S, Chua N (1998) The *Arabidopsis diminuto/dwarf1* gene encodes a protein involved in steroid synthesis. Plant Cell. 10:1677–1690.

Kozbial PZ, Mushegian AR (2005) Natural history of S-adenosylmethionine-binding proteins. BMC Struct Biol. 5:19.

Lai MH, Bard M, Pierson CA, Alexander JF, Goebl M, Carter GT, Kirsch DR (1994) The identification of a gene family in the Saccharomyces cerevisiae ergosterol biosynthesis pathway. Gene. 140:41-49.

Mathews FS (1991) New flavoenzymes. Curr Opin Struc Biol. 1:954-967.

Moore JT, Gaylor JL (1969) Isolation and purification of an S-adenosylmethionine: delta 24-sterol methyltransferase from yeast. J Biol Chem. 244(23):6334-6340.

Morikawa T, Mizatani M, Aoki N, Watanabe B, Saga H, Saito S, Qikawa A, Suzuki H, Sakurai N, Shibata D (2006a) Cytochrome P450 CYP710A encodes the sterol c-22 desaturase in Arabidopsis and tomato. Plant Cell. 18:1008-1022.

Morikawa T, Mizutani M, Ohta D (2006b) Cytochrome P450 subfamily CYP710A genes encode sterol c-22 desaturase in plants. Biochem Soc T. 34:1202-1205.

Nagano M, Uchimiya H, Kawai-Yamada M (2012) Plant sphingolipid fatty acid 2-hydroxylases have unique characters unlike their animal and fungus counterparts. Plant Signal Behav. 7(11):1388–1392.

Nebert DW, Gonzalez FJ (1987) P450 genes: structure, evolution, and regulation. Annu Rev Biochem. 56:945-993.

Nelson DR, Schuler MA, Paquette SM, Werck-reichart D, Bak S (2004) Comparative genomics of rice and *Arabidopsis*. Analysis of 727 cytochrome P450 gene and pseudogenes from a monocot and a dicot. Plant Physiol. 135:756-772.

Nishino H, Nakaya J, Nishi S, Kurosawa T, Ishibashi T (1997) Temperature induced differential kinetic properties between an initial burst and the following steady state in membrane bound enzymes: Studies on lathosteroid 5-Desaturase. Arch Biochem Biophy. 339(2):298-304.

Nithya RS, Anuja MM, Swathy SS, Rajamanickam C,Indira M (2011) Effects on spermatogenesis in swiss mice of a protein isolated from the roots of *Ricinus communis* (Linn.) (*Euphorbiaceae*). J Hazard Mater. 187(1–3):386–392.

Normen L, Dutta P, Lia A, Anderson H (2000) Soy sterol esters and β-sitosterol ester as inhibitors of cholesterol absorption in human small bowel. Am J Clin Nutr. 71:908-13.

Pata MO, Hannun YA, Ng CK-Y (2010) Plant sphingolipids: decoding the enigma of the Sphinx. New Phytol. 185:611-630 doi:10.1111/j.1469-8137.2009.03123.x.

Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer ELL, Eddy SR, Bateman A, Finn RD (2012) The Pfam protein families database. Nucleic Acids Res. 40: Database issue doi:10.1093/nar/gkr106.

Qian B, Raman S, Das R, Bradley P, McCoy AJ, Read RJ, Baker D (2007) High-resolution structure prediction and the crystallographic phase problem. Nature. 450:259–264.

Rahier A, Taton M, Bouvier-nave P, Schmitt P, Benveniste P, Schuber F, Narula AS, Cattel L, Anding C, Place P (1986) Design of high energy intermediate analogues to study sterol biosynthesis in higher plants. Lipids. 21:52-62.

Schaeffer A, Bronner R, Benveniste P and Schaller H (2001) The ratio of campesterol to sitosterol that modulates growth in *Arabidopsis* is controlled by sterol methyl transferase 2;1. Plant J. 25:605-615.

Schaller H, Bouvier- Nave P, Benveniste P (1998) Overexpression of an Arabidopsis cDNA encoding a sterol c-24 methyltransferase in tobacco modifies the ratio of 24-methyl cholesterol to sitosterol and is associated with growth reduction. Plant Physiol. 118:461-469.

Schaller H. (2004) New aspects of sterol biosynthesis in growth and development of higher plants. Plant Physiol. 42:465-76.

Schneider TD, Stephens RM (1990) Sequence Logos: A New Way to Display Consensus Sequences. Nucleic Acids Res. 18:6097-6100.

Schrick K, Mayer U, Martin G, Bellini C, Kuhnt C, Schmidt J, Jürgens G (2002) Interactions between sterol biosynthesis genes in embryonic development of Arabidopsis. Plant J. 31:61-73.

Schubert HL, Blumenthal RM, Cheng X (2003) Many paths to methyltransfer: a chronicle of convergence. Trends Biochem Sci. 28:329-335.

Senthil Kumar M, Wang K, Mysore KS (2013) *AtCYP710A1* gene-mediated stigmasterol production plays a role in imparting temperature stress tolerance in *Arabidopsis thaliana*. Plant Signal Behav. 8(2):e23142 doi: 10.4161/psb.23142

Sundaram P, Djerassi C (1997) A convenient synthesis of progesterone from stigmasterol. J Org Chem. 42(22):3633-3634.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA 5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance and maximum parsimony methods. Mol Biol Evol. 28(10):2731–2739.

Taton M, Rahier A (1991) Properties and structural requirements for substrate specificity of cytochrome P450-dependent obtusifoliol 14 α- demethylene from maize (*Zea mays*) seedlings. Biochem J. 277:483-492.

Tong Y, McCourt BS, Guo D, Mangla AT, Zhou WX, Jenkins MD, Zhou W, Lopez M, Nes, WD (1997) Steriochemical features of C- methylation on the path to delta24 (28)- methylene and delta24 (28)- ethylidene sterols: studies on the recombinant phytosterol methyl transferase from *Arabidopsis thaliana*. Tetrahedron Lett. 38:6115-6118.

Umebayashi K, Nakano A (2003) Ergosterol is required for targeting of tryptophan permease to the yeast plasma membrane. J Cell Biol. 161:1117–31.

Venkatramesh M, Guo DA, Jia Z, Nes WD (1996) Mechanism and structural requirements for transformation of substrates by the (S)- adenosyl- L- methionine: delta 24(25)- sterol methyl transferase from *Saccharomyces cerevisiae*. Biochim Biophys Acta. 1299(3):13-24.

Whitaker BD, Gapper NE (2008) Ripening –specific stigmasterol increase in tomato fruit is associated with increased sterol C-22 desaturase (CYP710A11) gene expression. Agric Food Chem. 56:3828-3835.

Wlodarski T, Kutner J, Towpik J, Knizewski L, Rychlewski L, Kudlicki A, Rowicka M, Dziembowski A, Ginalski K (2011) Comprehensive structural and substrate specificity classification of the *Saccharomyces cerevisiae* methyltransferome. PLoS One. 6(8):e23168 doi: 10.1371/journal.pone.0023168