Research in Plant Biology, 2(1): 41-49, 2012 ISSN : 2231-5101
www.resplantbiol.com

Research Article

Comparative analysis and partial functional annotation of phytosterol desaturase gene in plants of different botanical families

Raksha Bawankar[a], Pooja Singh[a], Rathinasamy Subashkumar[b], Govindasamy Vivekanandhan[b], Tha. Thayumanvan[c] and Subramanian Babu[a]

[a]School of Biosciences and Technology, VIT University, Vellore 632014, India
[b]PG and Research Department of Biotechnology, Kongunadu Arts and Science College, Coimbatore 641029, India
[c]School of Biotechnology, Dr.GRD College of Science, Coimbatore 641014, India

***In silico* analysis and comparison of the sterol desaturase gene from different plants species like Arabidopsis, rice, tomato and castor covering wide range of botanical families was done. The results indicate that Arabidopsis and tomato gene share 50% homology and rice-castor share 68% homology. The homology of sequences between Arabidopsis with other plants is negligible confirming the uniqueness of the gene in every botanical family of plants. Based on the translated amino acid sequences, the conserved domains in the protein were predicted by functional annotation including all enzymes in the database. Although the genes varied, the functional domains of Arabidopsis-tomato and rice-castor were similar. The major domain with specific conserved score was obtained for P450 super family proteins in the case of Arabidopsis and tomato. Rice and castor showed a conserved domain of sterol desaturase. An interior sequence of Arabidopsis and tomato showed a conserved domain for brassinosteroid oxidase protein. Based on the annotations, tomato is predicted to produce brassinosteroid as major sterol as like Arabidopsis, which belongs to Brassicaceae family. Rice and castor are predicted to use the desaturase gene for the production of stigmasterol. The results warrant further *in vitro* and *in planta* research to understand the functional diversity of the sterol desaturase gene in these plant species.**

Key Words: Cytochrome P450; Desaturase; Sitosterol; Stigmasterol

Phytosterol composition in plants is known for its high complexity compared with sterols of other eukaryotes. In contrast to the presence of single main sterol in many eukaryotes including yeasts, plant sterols usually occur as mixture of many sterols, each differing structurally. Sitosterol, stigmasterol, isofucosterol, campesterol and 24-methylcholestrol are some of the end sterols common in most plants (Schaller, 2003; Benveniste, 2004). In vascular plants, sitosterol and stigmasterol are the major sterols (Arnqvist et al., 2008). Altered ratio of sitosterol to stigmasterol affects the membrane properties upon stress. Morikawa et al. (2006) showed that the synthesis of stigmasterol is mediated by sterol C-22 desaturases. Thus, stigmasterol is known to be formed from sitosterol by a C-22 desaturation in the side chain. This relationship between sitosterol and stigmasterol biosynthetically has been a subject of some controversy. 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 this gene in plant species is still incomplete. As a preliminary step towards complete characterization of diversified sterol biosynthesis in plants, we attempted to compare the desaturase gene at the DNA sequence and annotated protein sequence level.

Materials and Methods

Comparison of nucleotide and amino acid sequences

With the objective to understand the sterol desaturase gene distribution and sequence wide variation among plant species, the nucleotide sequences of *Arabidopsis thaliana*, rice (*Oryza sativa*), tomato (*Solanum lycopersicon*) and castor (*Ricinus communis*) belonging to families Brassicaceae (Cruciferae), Poaceae (Graminnae), Solanaceae and Euphorbiaceae respectively, were obtained from the NCBI eukaryote genome data base.

ClustalW 2.0.12 multiple sequence alignment program (http://www.ebi.ac.uk/clustalw) was used to align the sterol desaturase gene sequences. The percentage homology between the genes of these source plants were also determined using the same software and documented. The amino acid sequences of the desaturase protein was also aligned to infer the conserved and deviating regions in the sequences among different plants.

Comparison of conserved domains in proteins

Conserved Domain Search Service of NCBI was used for comparing the conserved domains in the protein. The tool identifies the conserved domains present in a protein sequence by using RPS-BLAST (Reverse Position-Specific BLAST) to compare a query sequence against position-specific score matrices that have been prepared from conserved domain alignments present in the Conserved Domain Database (CDD).

To understand the diversified sterol biosynthesis in Arabidopsis with other plants like rice, protein domain architecture was determined based on functional annotation of the amino acid sequences using the Conserved Domain Architecture Retrieval Tool (CDART). CDART performs similarity searches of the NCBI Entrez Protein Database based on domain architecture, defined as the sequential order of conserved domains in proteins. The algorithm finds protein similarities across significant evolutionary distances using sensitive protein domain profiles rather than by direct sequence similarity. Proteins similar to a query protein are grouped and scored by architecture.

Results and Discussion

Phytosterols, as like cholesterol in mammals and ergosterol in fungi, are involved in the formation of membrane microdomains that serve as a platform for crucial physiological processes such as signal transduction, vesicular trafficking and cytoskeleton organization (Mongrand et al., 2004; Borner et al., 2005). They are also involved in biosynthesis of steroid hormones (Schaller, 2003). Since sitosterol and stigmasterol are the major phytosterols in vascular plants and stigmasterol is produced at the expense of sitosterol, the desaturase enzyme involved in between these two sterols is a candidate target for modulations of sterol biosynthesis in plants. The sterol desaturase proteins belong to a small cytochrome P450 subfamily having four members, denoted by CYP710A1-A4. Transgenic Arabidopsis expressing A1 and

A4 contained increased level of free stigmasterol and concomitant decrease in the level of sitosterol. Our sequence analysis in the present study was performed on A4 protein coding desaturase gene.

Sterol desaturase gene sequences from wide range of plants were retrieved as described and the sequences were compared. The results of multiple sequence alignment is presented in the Fig. 1. The percentage homology of sequences are presented in Table 1. Size of the gene in Arabidopsis, rice, tomato and castor are 1.928, 1.252, 1.776 and 0.831 kb respectively. The sequence homology exhibited by Arabidopsis gene with that of rice gene was 2% and with tomato and castor, 50% and 4% respectively. The sequence homology was 5% between rice and tomato, 68% between rice and castor and 3% between tomato and castor (Table 1). Translated amino acid sequences of these genes were aligned using ClustalW program as described in the materials and methods. The results of the alignment are presented in Fig 2.

Fig 1. Alignment of sterol desaturase gene sequences of different plant species

```
RICE             ------------------------------------------------------------
RICINUS          ------------------------------------------------------------
ARABIDOPSIS      ATACCATTGCAATTTGCAAACTGCTTCAATCCAATTTCCCTAAACCTCTTAGAAAATTCT 60
TOMATO           ------------------------------------------------------------

RICE             ------------------------------------------------------------
RICINUS          ------------------------------------------------------------
ARABIDOPSIS      GCAAAACCCACCAAACAAAGAACACCAAGAAAGACAAAAGGCAGGCAAGACTACAAAGCA 120
TOMATO           ------------------------------------------------------------

RICE             ------------------------------------------------------------
RICINUS          ------------------------------------------------------------
ARABIDOPSIS      AAACCTTCTTCTTCTTCTTCTCGTTACTCAAATCATTTCCATATATACGCAAAAACC 180
TOMATO           -------------------------------------------GAAGAAGAAGAAGAAGA 17

RICE             ------------------------------------------------------------
RICINUS          ------------------------------------------------------------
ARABIDOPSIS      ATTAAGACAAGCTA---GAAACAAAGAAACATGGTTTTCTCTGTTTCTATATTTGCCTCT 237
TOMATO           AGAGAGCTATATTGTACGGTACTTCATCACATGGCATCCATTTGGGGTTTGTTATCTCCA 77

RICE             ------------------------------------------------------------
RICINUS          ------------------------------------------------------------
ARABIDOPSIS      CTTGCACCATACTTAATCTCAGCATTTCTTCTTTTCCTTCTCGTCGAGCAACTCTCTTAC 297
TOMATO           TGGATACCTTATTTCATTTCTTTCATAGCTTTTTTACTTCTTCTTGAACAGATCTCTTAC 137

RICE             ------------------------------------------------------------
RICINUS          ------------------------------------------------------------
ARABIDOPSIS      CTTTTCAAGAAACGAAACATCCCTGGCCCTTTCTTCGTCCCTCCAATCATCGGAAACGCC 357
TOMATO           ATCAAGAAGAAGCGTTTTCTTCCTGGCCCAACTCTTGTATTCCCCTTCCTTGGCAACGTA 197

RICE             ------------------------------------------------------------
RICINUS          ------------------------------------------------------------
ARABIDOPSIS      GTTGCACTCGTTCGTGACCCCACTTCCTTCTGGGACAAGCAATCCTCCACGGCGAA---- 413
TOMATO           ATTCCCTTAGTCACAAATCCAACTAAATTCTGGGACCTTCAATCAGCTTTAGCTAAGTCT 257

RICE             --------------------------------------------------ACTCTCTCTC 10
RICINUS          ------------------------------------------------------------
ARABIDOPSIS      --CATCTCAGGCCTCTCCGCCAACTACCTCATCGGAAAATTCATCGTCTATATCAGAGAC 471
TOMATO           ACTAGCCATGGTTTTTCTGTTAACTACATCATAGGTAAGTTCATTCTTTACATCCACTCA 317

RICE             TCCACGCTTTACAACTCAACCCCCGCACCACCACCAGTCCACCACCACAATCCCAGGGCG 70
RICINUS          ------------------------------------------------------------
ARABIDOPSIS      ACTGAGCTTTCCCATCAAATCTTCTCAAACGTTCGTCCCGATGCCTTCCACCTTATCGGA 531
TOMATO           ACTGACCTCTCTCATAAGGTCTTTGCCAATGTCCGCCCTGACGCTTTCCATCTTATCGGT 377

RICE             GCGCCGCGGCCGGCAATGGCGGGCGGCGGCGGCGAGTACCTGCGCCAGTTCGTCGAGGAG 130
RICINUS          ---------------ATGGAGGTGGTGGAGAACAAGCAGTTGCAGTATTTCTTGGAAGAG 45
ARABIDOPSIS      CATCCCTTTGGTAAGAAGCTCTTCGGTGACCACAACCTTATATACATGTTCGGCGAGGAT 591
TOMATO           CACCCTTTTGGGAAAAAGCTATTCGGCGAACATAACTTGATTTACATGTTTGGGCAAGAA 437
                                * *      *  *      *     *       **     * **
RICE             -ACGGCCTGGTACAACGAGATCTTCCTCAGCCATGTGGTCCCGGGCGACTGGTGGCGCGC 189
RICINUS          -ACAGCGATGTATAACAACATTGTCTTAGGACATCTATTGCCTTCTGGATGGTGGGCCAC 104
ARABIDOPSIS      CACAAATCCGTTCGCCGTCAGCTAGCACCTAACTTCACTCCCAAGGCACTGTCCACTTAC 651
TOMATO           CATAAAGACCTTCGCCGACGAATTGCCCCAAATTTTACCCCTAAAGCTCTGGGAACTTAC 497
                  *        *    *                 *      *        **        *
```

```
RICE             CCTCCCCCACCCGCTCCAGTCGTGGCTCCGCAACGGCCTCGGCGGCTACCTCATCTACTT 249
RICINUS          GCTTCCCAGGTTCTTGCAGACGTGGCTTCGTAACTACATTGCTGGAACTCTCCTTTACTT 164
ARABIDOPSIS      TCTGCCCTCCAGCAATTAGTTATTCTCCGTCATCTACGGCAGTGGGAGGGGAGTACCTCC 711
TOMATO           ACTGATATTCAACAGAGGATTATTATCAAACACTTCAAGTCCTGGTTAGATGAAGCATCC 557
                  **                   *        *           **

RICE             CGCCTGCGGCTTCCTCTGGTGCTTCGTCATCTACTACTGGAAGCGCCACGCCTACAT-CC 308
RICINUS          CATCTCTGGCTTCCTGTGGTGTTTCTACATTTACTACTTGAAACGCAACGTTTATCT-TC 223
ARABIDOPSIS      GGCGGATCTCGT---CCAGTGTCACTGCGACAGCTTGTCCGTGAACTCAACCTAGAAACT 768
TOMATO           AAATCCCCTAACACCCCAATCCCGCTTCGTCTACTTTGCAGGGATATGAACTTGGATACT 617
                 *    *  *     **                 *

RICE             CCAAAGATTCTATACCTACAATCGAAGCT--ATGAAGAAGCAAATAATTGTTGCATCAAA 366
RICINUS          CTAAAGATGCTATCCCTACAACAAAAGCC--ATGCTCTTGCAAATATATGTTGCTATGAA 281
ARABIDOPSIS      TCGCAGACGGTTTTTGTTGGACCCTACCTTGACAAGGAAGCCAAGAACAGGTTCCGTACT 828
TOMATO           TCTCAGACTGTGTTCGTTGGTCCATACTTGGATGGAGAATCGAGAAAGAGATTTAATGTT 677
                     ***   * *   *        *     *        * *  *   * *

RICE             GGCTATGCCTCTCT--ATTGTGCCCTTCCAACCTTATCTGAGTACATGGTTGAGAATGGA 424
RICINUS          GGCTATGCCATGGT--ACTGTGCTCTTCCATCTCTTTCTGAGTACATGGTTGAAAATGGC 339
ARABIDOPSIS      GATTATAATTTGTTCAATCTCGGATCCATGGCTCTCCCCATAGACCTCCCTGGCTTTGCG 888
TOMATO           GATTACAATTACTTCAATGTTGGGTTAAGGAAACTTCCTGTTGATTTACCGGGTTTTGCC 737
                   *  **        *  *    *            *  *     *  *    *    **

RICE             TGGACACAGTGTTATGTTAATATCAGTGAAGTTGGTTGGCCAATGTACCTGGTTTATCTG 484
RICINUS          TGGACTAAGTGTTTCTCATACATATCTGATGTTGGCTGTTTTGGCTACATCATATATTTA 399
ARABIDOPSIS      TTCG-GCGAGGCTCGCAGGGCTGTAAAGAGGCTAGGGGAG--ACGCTCGGCATTTGCGCG 945
TOMATO           TTCA-GAAATGCTAGATTAGCAGTTGGGAGATTAGTTGAC--ACCCTTTCGGTTTGTGTG 794
                 *         * *              **   * *  *              * *

RICE             GCTTTATATCTTATCTTTGTTGAGTTTGGAATTTA--CTGGATGCACAGAGAGTTGCATG 542
RICINUS          ATTGCGTATCTTGTTATGGTGGAGTTTGGGATTTA--CTGGATGCACAGAGAATTGCATG 457
ARABIDOPSIS      GGAAAATCCAAAGCGAGGATGGCAGCAGGAGAAGAACCAGCATGCCTAATCGATTTCTGG 1005
TOMATO           GAACAAAGCTTAAACAAGATGAAAAACGAAGAAGAACCCACATGCTTGATTGATTTCTGG 854
                                    *       *      *  *   ****        ** *  *

RICE             ACATAAAGCCATTGTACAAGTACCTGC--------------------ACACATACCACCA 582
RICINUS          ATATAAAACCTTTATACAAGTATCTTC--------------------ATGCAACCCATCA 497
ARABIDOPSIS      ATGCAGGCGATCGTCGCGGAGAAT-----------------------CCGCAGCCACCG 1041
TOMATO           ATGCAGGAAAATTTAAGAGAGATTAACGAAGCTAAGATCAATGGATTACAAAAGCCATTT 914
                 *   *                ***

RICE             TATTTACAACAAGGAGAATACCCTATCACCATTTGCAGGACTAGCATTCCATCCACTGGA 642
RICINUS          CATCTATAACAAGCAGAATACTCTTTCTCCTTTTGCCGGTTTGGCATTTCATCCGCTTGA 557
ARABIDOPSIS      CA-CTCGGGAGACGAAGAGATCGGTGGTTTGCTCTTCGATTTTCTCTTTGCCGCTCAAGA 1100
TOMATO           CA-GTACAGTAACAAGGAACTTGGAGGTTACCTGTTCGACTTCCTCTTTGCTGCTCAAGA 973
                  *  *      *  *  *              *    *   *    **     * *  **

RICE             TGGGATTTTGCAAGCCATACCGCATGTGTTTGCGCTCTACCTTATCCCAACACAC-TTCA 701
RICINUS          CGGAATACTCCAAGCAGTCCCACATGTTATTGCTCTCTTCATTGTACCTACACAT-TTCA 616
ARABIDOPSIS      CGCGTCCACGTCATCACTCCTCTGGGCCGTGACGCTTCTTGATTCAGAGCCAGAGGTGCT 1160
TOMATO           TGCTTCTACTTCTGCTCTGTTATGGGCAATCGTGCTTCTAGATTCTCACCCACAAGTTCT 1033
                  *            *  *       *   *    **      *       ** *  * *

RICE             GGACACACATTGCTCTCTTGTTCATAGA----GGCCGTGTGGACAACTAACATCCATGAC 757
RICINUS          GGACACACATAGGTCTCTTATTCCTAGA----GGCCATATGGACCGCAAACATTCATGAT 672
ARABIDOPSIS      GAACAGAGTAAGAGAAGAAGTAGCCAAGATCTGGTCACCTGAGTCCAACGCCTTGATCAC 1220
TOMATO           GGAGAAAGTTCGGTCGGATGTAGCGAGATTCTGGTCGCCAGAATCTGAGGAGCCGCTGAC 1093
                 * * * *    *        *    *      ** *    *               * *

RICE             TGCATTCACGGCAAGGTTTGGCCGGTCATGGGTGCTGGCTATCACACCAT-TCACCATAC 816
RICINUS          TGCATACATGGTAAGCTATGGCCGGTGATGGGTGCTGGCTATCACACCAT-CCACCATAC 731
ARABIDOPSIS      CGT-TGATCAGCTCGCAGAGATGAAGTATACACGCTCCGTGGCGCGTGAGGTCATTAGAT 1279
TOMATO           GGC-GGAAATGCTCACGGAAATGAAGTACCTGGAAGCGGTGGCGCGTGAGATAATCAGAA 1152
                  *        *                *           *  * *   *    *  * *

RICE             AACATACCGTCACAACTATGGCCAC-TACACCGTGTGGATGGACTGGATGTT--CGGCAC 873
RICINUS          TACATACCGCCACAATTATGGTCAT-TATACCATATGGATGGATTGGATGCT--AGGAAC 788
ARABIDOPSIS      ACAGGCCTCCTGCAACTATGGTCCCACACGTCGCTGCTATAGACTTCCCTCTCACGGAAA 1339
TOMATO           TCAGAGCTCCGGCGACAATGGTGCCACATATTGCCGGCGAAGAATTCCGGTTAACCGAAG 1212
                       *     * *  ****      *             ** *      *    * *

RICE             CCTTCGAGAGCC---AGAAGA----TATCTTGAAGAAGGATTAGGATCGAATTGTTAAGC 926
RICINUS          CCTTCGTGACCCTGAGGAAGA----TGCATTCGAGAAGGTGAAGTGA------------- 831
ARABIDOPSIS      CGTACACTATCCCAAAAGGTACAATTGTCTTTCCTTCGGTTTTCGACTCCTCGTTCCAAG 1399
TOMATO           ATTACGTTATCCCAAAAGGAACAATTGTGTTCCCGTCGGTTTTTGATTCATCATTTCAGG 1272
                   * *   * **        *    *   **      **
```

```
RICE             ACGGCGCGACGTTCGCCTCTCGTCTTTGTAGCAGTCGGAAGAAGGGGGTTCTATGTATTT 986
RICINUS          ------------------------------------------------------------
ARABIDOPSIS      GGTTTACTGAACCGGACCGGTTTGATCCTGACCGGTTTAGCGAGACAAGACAAGAGGACC 1459
TOMATO           GTTTTCCTGAACCGGAGAAATTTGAACCGGACCGGTTCATGGAGGAGAGACAAGAGGAGC 1332

RICE             ATCCTGCTCGATGCTATCTGTTATCTTTCATGTACCAATGTGCTGTTTCAGTATGCTAGG 1046
RICINUS          ------------------------------------------------------------
ARABIDOPSIS      AGGTGTTCAAACGCAACTTCCTAGCTTTTGGATGGGGGCCTCACCAATGCGTAGGCCAGC 1519
TOMATO           GGGTTTACAAAAAGAACTTTCTAGCATTAGGTGCTGGGCCCCATGCGTGTGTCGGACAGA 1392

RICE             TGAACATAAAAGTGAACTT-----TAGTGAGTTCATCATCCCTGGAATGCAATGGAATCG 1101
RICINUS          ------------------------------------------------------------
ARABIDOPSIS      GTTACGCGTTGAACCACCTCGTACTCTTCATTGCAATGTTCTCGTCGTTGTTGGATTTCA 1579
TOMATO           AGTATGCTATTAACCACTTGATGCTTATCATTGCTATGTTTACGGCTCTGATTGATTTCA 1452

RICE             AGAGATATATCCTGGGTAGGTTCCAGGTGTTTGTACCATTCCTGAATTTCTCTGTACTTC 1161
RICINUS          ------------------------------------------------------------
ARABIDOPSIS      AGAGGCTTCGATCAGACGGTTGTGATGAGATCGTGTACTGCCCTACTATATCGCCCAAGG 1639
TOMATO           AGAGACACAAAACCGACGGCTGCGATGACATCTCGTATATTCCAACCATTGCTCCAAAGG 1512

RICE             TTTGTGTCAAAGAACATGATGAACAGTTGTGGCCAGCTTTTGTACAGCAATAACTGTTTT 1221
RICINUS          ------------------------------------------------------------
ARABIDOPSIS      ATGGGTGCACCGTGTTC-TTGTCTAGGCGCGT-CGCAAAGTATCCCAACTTTTCCTAAAT 1697
TOMATO           ATGATTGCAAAGTTTTCCTTGCACACAGGTGCACACGATGATTGCCCATTTTTTAAAAAA 1572

RICE             ATTGTAACTGCTGCCATCCTTTTGTATCATG----------------------------- 1252
RICINUS          ------------------------------------------------------------
ARABIDOPSIS      TTTTTGGATTTTTCTTAAATTTATTCTTGGTGGTCAAAGAAACCAGTGTGAATTTATCAC 1757
TOMATO           ATAATCAATTTGGCCCCATTTTAATTATTTTT-CTATTTTTTTCCTTGTTTTTTTTTTAA 1631

RICE             ------------------------------------------------------------
RICINUS          ------------------------------------------------------------
ARABIDOPSIS      TTTTCGTGATGTGTCTCTCGAAGTGAGTGGCCCA--CGAGTGTTCTTGTCTTATTGAATT 1815
TOMATO           TTTTAAATTTTAGTGTGTGAATACACGTAACTTGATCAAATGATCAAGATTC------TT 1685

RICE             ------------------------------------------------------------
RICINUS          ------------------------------------------------------------
ARABIDOPSIS      TGAGAAGACTTGTCTTTGTTTTGCCATGTCTTCACCA-TATTTGTT-TTGTGTTGTCTTC 1873
TOMATO           TTAAAATGTTTGTTGGAGCAAT-TGATGGAGTGACAAATATTTGGGATTGTGTGGAAGAT 1744

RICE             -------------------------------------------------------
RICINUS          -------------------------------------------------------
ARABIDOPSIS      CAGTGAACATCAGTGATGTTACAACGTGTTTCATACTTGTATTTTTTTTCTTGCC 1928
TOMATO           TAATAAAAGAAATTCCTTTTA-AGGGCCCTAAA---------------------- 1776
```

Table 1. Homology comparison of sterol desaturase gene sequences among different plant species

Seq A	Name	Length (nt)	Seq B	Name	Length (nt)	Homology Score
1	Arabidopsis	1928	2	Rice	1252	2
1	Arabidopsis	1928	3	Tomato	1776	50
1	Arabidopsis	1928	4	Ricinus	831	4
2	Rice	1252	3	Tomato	1776	5
2	Rice	1252	4	Ricinus	831	68
3	Tomato	1776	4	Ricinus	831	3

To further understand the desaturase protein at the active domain level, prediction and functional annotation of active domains in the amino acid sequences was performed using the Conserved Domain Search Service of NCBI as described and the results are presented in Fig. 3.

The results indicate that the Arabidopsis and tomato desaturase protein sequences share a common domain of cytochrome P450 superfamily proteins. Rice and castor shared the common domain of sterol desaturase (ERG3). Other than the P450 subfamily domains and non-specific domains, the major annotated domains of the desaturase protein from these representative plant species are listed in Table 2. It is found that Arabidopsis and the tomato desaturase proteins do not have the sterol desaturase domain that is present in rice and castor. The active domains in the Arabidopsis desaturase amino acid sequences that correlate with the protein domain database of plants include cytochrome P450 superfamily and subfamily, ent-kaurenoic acid hydroxylase involved in the diterpenoic steviol glycoside biosynthesis, abscisic acid dehydroxylase, brassinosteroid oxidase, carotene beta ring hydroxylase and steroid alpha hydroxylase. The active domains in the tomato desaturase amino acid sequences that correlate with the protein domain database of plants include the same as like Arabidopsis. ERG3 sterol desaturase, fatty acid and carotene hydroxylase, fatty acid hydroxylase superfamily are the matched domains in the rice and castor desaturase amino acid sequences.

Table 2. Conserved domains of the desaturase predicted proteins of different plant species

S.No	Plant species	Amino acids in the gene	Amino acids range of the annotated domain	Conserved domain ID	Function
1	*Arabidopsis thaliana*	495	75-450	Cytochrome P450 superfamily	Degradation of environmental toxins and mutagens
			105-490	Cytochrome P450 (cypX)	Secondary metabolite synthesis and transport
			105-495	Ent-kaurenoic acid hydroxylase	Glycoside biosynthesis
			100-490	Abscisic acid dehydroxylase	Signal transduction
			100-480	Brassinosteroid 6-oxidase	Brassinosteroid biosynthesis
			285-490	Carotene beta ring hydroxylase	Carotenoid biosynthesis
			195-440	Steroid 22 alpha hydroxylase	Steroid biosynthesis
2	*Solanum lycopersicum*	501	40-460	Cytochrome P450 superfamily	Degradation of toxins and mutagens
			110-500	Cytochrome P450 subfamily	Secondary metabolite synthesis and transport
			30-500	Ent-kaurenoic acid hydroxylase	Glycoside biosynthesis
			15-500	Abscisic acid hydroxylase	Signal transduction
			15-450	Steroid 22 alpha hydroxylase	Steroid biosynthesis
			15-450	Brassinosteroid 6-oxidase	Brassinosteroid biosynthesis
3	*Oryza sativa*	274	40-270	ERG3 sterol desaturase	Sterol biosynthesis
			130-240	Fatty acid and carotene hydroxylase	Zeazanthin biosynthesis
			40-270	Fatty acid hydroxylase superfamily	Carotenoid biosynthesis
4	*Ricinus communis*	276	35-276	ERG3 sterol desaturase	Sterol biosynthesis
			140-240	Fatty acid and carotene hydroxylase	Zeazanthin biosynthesis
			35-276	Beta carotene hydroxylase	Carotenoid biosynthesis

The results of these *in silico* analyses indicate that the gene sequences are totally unique to each species. The size of the gene also varied among the plants chosen for the study. This represents the differential function and regulation of the gene in sterol biosynthesis among different plant species. The homology score of the nucleotide sequence is 50% between *Arabidopsis* and tomato and 68% between rice and castor. The comparison of sequences between *Arabidopsis* and rice, as well as tomato –rice and castor-rice showed very minimum homology which is negligible. This indicates that among the sterol desaturase genes of plants, there are desaturase families corresponding to botanical species. It is also interesting that Arabidopsis do not share complete homology with any of the desaturase gene sequences available in the database.

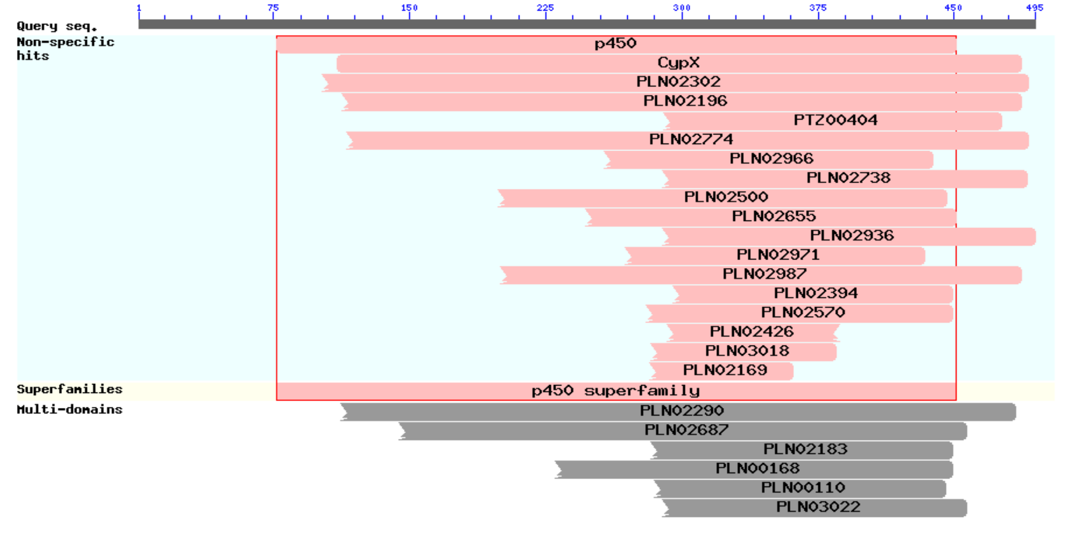
Fig 2. Alignment of sterol desaturase amino acid sequences of different plant species

```
Rice             ------------------------------------------MAGGGGEYLRQFVEETAW 18
Ricinus          ------------------------------------------MEVVENKQLQYFLEETAM 18
Arabidopsis      MVFSVSIFASLAPYLISAFLLFLLVEQLSYLFKKRNIPGPFFVPPIIGNAVALVRDPTSF 60
Tomato           MASIWGLLSPWIPYFISFIAFLLLLEQISYIKKKRFLPGPTLVFPFLGNVIPLVTNPTKF 60
                                                           :    .: :  . : *

Rice             YN---------------EIFLSHVVPGDWWRALPHPLQSWLR-------NGLGGYLIYFA 56
Ricinus          YN---------------NIVLGHLLPSGWWATLPRFLQTWLR-------NYIAGTLLYFI 56
Arabidopsis      WDKQSSTANIS--GLSANYLIGKFIVYIRDTELSHQIFSNVRPDAFHLIGHPFGKKLFGD 118
Tomato           WDLQSALAKSTSHGFSVNYIIGKFILYIHSTDLSHKVFANVRPDAFHLIGHPFGKKLFGE 120
                 ::               : .:.:.:       *.: : : :*       .   *  ::

Rice             CGFLWCFVIYYWKRHAYIPKDSIP------------------------------------ 80
Ricinus          SGFLWCFYIYYLKRNVYLPKDAIP------------------------------------ 80
Arabidopsis      HNLIYMFGEDHKSVRRQLAPNFTPKALSTYSALQQLVILRHLRQWEGSTSGG-SRPVSLR 177
Tomato           HNLIYMFGQEHKDLRRRIAPNFTPKALGTYTDIQQRIIIKHFKSWLDEASKSPNTPIPLR 180
                  .::: *   : . .  :. :  *

Rice             ---------TIEAMKKQIIVASKAMPLY-------------------------------- 99
Ricinus          ---------TTKAMLLQIYVAMKAMPWY-------------------------------- 99
Arabidopsis      QLVRELNLETSQTVFVGPYLDKEAKNRFRTDYNLFNLGSMALPIDLPGFAFGEARRAVKR 237
Tomato           LLCRDMNLDTSQTVFVGPYLDGESRKRFNVDYNYFNVGLRKLPVDLPGFAFRNARLAVGR 240
                          * :::     :  ::   :

Rice             -------CALPTLSEYMVEN-------------------------------------GWT 115
Ricinus          -------CALPSLSEYMVEN-------------------------------------GWT 115
Arabidopsis      LGETLGICAGKSKARMAAGEEPACLIDFWMQAIVAEN--------PQPPHSGDEEIGGLL 289
Tomato           LVDTLSVCVEQSLNKMKNEEEPTCLIDFWMQENLREINEAKINGLQKPFQYSNKELGGYL 300
                        *.  :  .    :                                     *

Rice             QCYVNISEVGWPMYLVYLALYL---------IFVEFGIYWMHR--------ELHDIKPLY 158
Ricinus          KCFSYISDVGCFGYIIYLIAYL---------VMVEFGIYWMHR--------ELHDIKPLY 158
Arabidopsis      FDFLFAAQDASTSSLLWAVTLLDSEPEVLNRVREEVAKIWSPESNALITVDQLAEMKYTR 349
Tomato           FDFLFAAQDASTSALLWAIVLLDSHPQVLEKVRSDVARFWSPESEEPLTAEMLTEMKYLE 360
                   :   :: .    :::    *         :  :..  *  .         * ::*

Rice             KYLHTYHHIYNKENTLSPFAGLAFHPLDGILQAIPHVFALYLIPTHFR--THIALLFIEA 216
Ricinus          KYLHATHHIYNKQNTLSPFAGLAFHPLDGILQAVPHVIALFIVPTHFR--THIGLLFLEA 216
Arabidopsis      SVAREVIRYRPPATMVPHVAAIDFPLTETYTIPKGTIVFPSVFDSSFQGFTEPDRFDPDR 409
Tomato           AVAREIIRIRAPATMVPHIAGEEFRLTEDYVIPKGTIVFPSVFDSSFQGFPEPEKFEPDR 420
                    :   :     . :. .*.  *   :    .   :.   :. : *:  ..   :  :

Rice             VWTTNIHDCIHGKVWPVMGAGYHTIHHTTYRHNYGHYTVWMDWMFGTLREPE-DILKKD- 274
Ricinus          IWTANIHDCIHGKLWPVMGAGYHTIHHTTYRHNYGHYTIWMDWMLGTLRDPEEDAFEKVK 276
Arabidopsis      FSETRQEDQVFKRNFLAFGWGPHQCVGQRYALNHLVLFIAMFSSLLDFKRLRSDGCDEIV 469
Tomato           FMEERQEERVYKKNFLALGAGPHACVGQKYAINHLMLIIAMFTALIDFKRHKTDGCDDIS 480
                 .   . .: :. : : .:* * *      * *:     : *   :  ::  . *  ..

Rice             --------------------------
Ricinus          --------------------------
Arabidopsis      YCPTISPKDGCTVFLSRRVAKYPNFS 495
Tomato           YIPTIAPKDDCKVFLAHRCTR----- 501
```

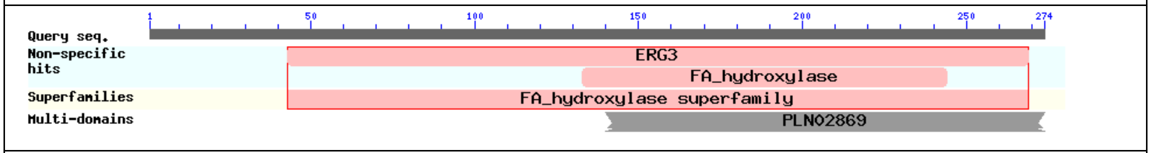
Morikawa et al. (2009) over-expressed the sterol C-22 desaturase gene in Arabidopsis and tomato and demonstrated that this CYP710A2 is responsible for stigmasterol production. Our results on the protein active domain level represents the sequence level justification of how *Arabidopsis* gene ectopically expressed in tomato could produce more stigmasterol.

Fig 3. Conserved domain comparison of desaturase protein in different plant species

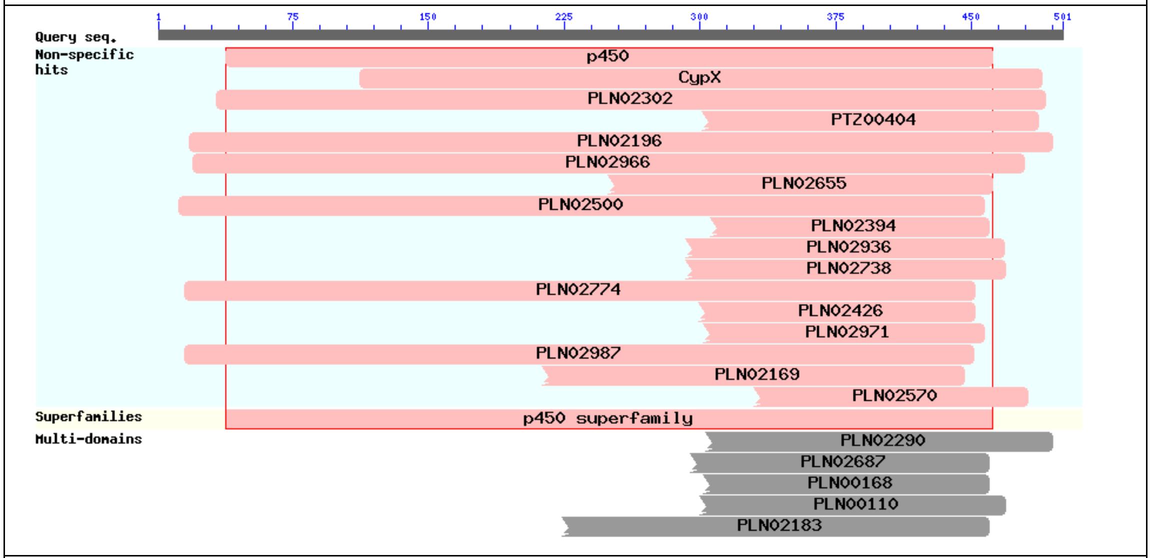
A. *Arabidopsis thaliana*



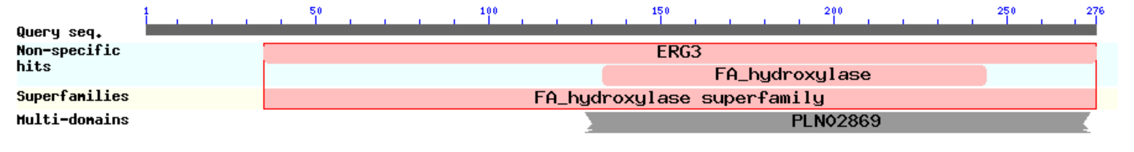
B. *Oryza sativa*



C. *Solanum lycopersicum*



D. *Ricinus communis*



In our study, sterol desaturase active conserved domain was obtained as the first specific hit in the rice and castor. Whereas the sequences resulted in the prediction of P450 family member proteins in *Arabidopsis* and tomato. Morikawa et al. (2009) demonstrated that the substrate specificities of Arabidopsis and tomato CYP710A1 are strict towards sitosterol and the A2 is capable of producing both stigmasterol and brassinosterol from sitosterol and campesterol, respectively. Based on our present functional annotations of the sequences, we suggest that the CYP710A4 could also have substrate specificity towards sitosterol. In addition, among the conserved domains with sterol biosynthesis, *Arabidopsis* and tomato resulted in a conserved domain for brassinosteroid oxidase also. *Arabidopsis* belonging to the Brassicaceae family is known to synthesise brassinosteroid as the major steroid compound. But the observation of brassinosteroid oxidase domain in the sterol desaturase gene of tomato, made in the study, represents the first report that the major sterols in tomato could be brassinosterol. This warrants further phylogenetic analysis of the sterol desaturase gene in the context of taxonomical nearness of Brassicaceae and Solanaceae family plants.

References

Arnqvist L., Persson, M., Jonsson, L., Dutta, P.C., Sitbon, F. 2008. Over-expression of CYP710A1 and CYP710A4 in transgenic Arabidopsis plants increases the level of stigmasterol at the expense of sitosterol. *Planta,* **227**: 309-317.

Benveniste, P. 2004. Biosynthesis and accumulation of sterols. *Annu. Rev. Plant Biol.,* **55**: 429-457.

Borner, G.H., Sherrier, D.J., Weimar, T., Michaelson, L.V., Hawkins, N.D., Macaskill, A., Napier, J.A., Beale, M.H., Lilley, K.S., Dupree, P. 2005. Analysis of detergent resistant membranes in Arabidopsis. Evidence for plasma membrane lipid rafts. *Plant Physiol.,* **137**: 104-116.

Mongrand, S., Morel, j., Laroche, j., Claverol, S., Carde, J.P., Hartmann, M.A., Bonneu, M., Simon-Plas, F., Lessire, R., Bessoule, J.J. 2004. Lipid rafts in higher plant cells: Purification and characterization of Triton X-100-insoluble microdomains from tobacco plasma membrane. *J. Biol. Chem.,* **279**: 36277-36286.

Morikawa T., Mizutani, M., Aoki, N., Watanabe, B., Saga, H., Saito, S., Oikawa, A., Suzuki, H., Sakurai, N., Shibata, D., Wadano, A., Sakata, K., Ohta, D. 2006. Cytochrome P450 CYP710A encodes the sterol C-22 desaturase in Arabidopsis and tomato. *Plant Cell*, **18**: 1008-1022.

Morikawa, T., Saga, H., Hashizume, H., Ohta, D. 2009. CYP710A gene encoding sterol C-22 desaturase in Physcomitrella patens as molecular evidence for the evolutionary conservation of a sterol biosynthetic pathway in plants. *Planta,* **229**: 1311-1322.

Schaller, H. 2003. The role of sterols in plant growth and development. *Progr. Lipid Res.,* **42**: 163-175.