American Journal of Biochemistry and Biotechnology

Original Research Paper

In silico Analysis of 4CL Family in *Scutellaria baicalensis* through Biocomputational Tools and Servers

[1,2†]Guoming Li, [3†]Xiaozhong Lan, [1,4]Xiaorong Shui, [1,2]Shian Huang, [1,2,5]Can Chen and [1,2,5]Wei Lei

[1]*Zhanjiang Key Laboratory of Biomedicine and Cardiovascular Disease Prevention in West Guangdong, Guangdong Medical University, Zhanjiang 524001, China*
[2]*Cardiovascular Medicine Center, Affiliated Hospital of Guangdong Medical University, Zhanjiang 524001, China*
[3]*Tibetan Collaborative Innovation Center of Agricultural and Animal Husbandry Resources, Tibet Agricultural and Animal Husbandry College, Nyingchi 860001, China*
[4]*Laboratory of Vascular Surgery, Guangdong Medical University, Zhanjiang 524001, China*
[5]*Laboratory of Cardiovascular Diseases, Guangdong Medical University, Zhanjiang 524001, China*

Article history
Received: 08-09-2016
Revised: 02-01-2017
Accepted: 26-01-2017

Corresponding Author:
Wei Lei and Can Chen
Laboratory of Cardiovascular Diseases, Guangdong Medical University, Zhanjiang 524000, China
Tel: 86-0759-2369147,
Fax: 86-0759-2231754
Email: leiwei2006@126.com
chencan_21@163.com

† These authors contributed equally to this work.

Abstract: The gene sequences of 4-coumarate:coenzyme A ligase (4CL) family of *Scutellaria baicalensis* came from the GenBank Database. With help of some bioinformatics tools such as Vector NTI Suite 8, ProtParam and SWISS-MODEL and so on, a series of biological information of their nucleic acid sequences and amino acid sequences were predicted and analyzed and the results were revealed as following: Two *4cl* member genes share high similarity in structure and properties on level of nucleic acid and amino acid molecular. 4CLs are hydrophic proteins without any transmembrane topological structure and some crucial motifs were found. The secondary structures of Sb4CLs are mainly composed of random coil and α-helix and the models of their tertiary structures were built. *In silico* analysis of Sv4CL was finished, which would pave for further studies of physichemical properties of 4CL family and its related molecular mechanism of flavonoid metabolic regulation.

Keywords: *Scutellaria baicalensis*, Flavonoid, Metabolic Regulation, 4-Coumarate:coenzyme A Ligase, Bioinformatics

Introduction

Bioinformatics is a science, which uses data information based on mathematics and computer science to understand biology. In the post genome era, researches of the protein structures and functions are the focus issues of molecular biology field and today, a number of computational software's and online servers are rapidly developed for identification and characterization of proteins and their encoded nucleotide acid sequences (Sivakumar *et al*., 2007; Lei *et al*., 2009). The physiochemical properties and biological function of the proteins can be well studied with bioinformatics methods (Ling *et al*., 2007; Lei *et al*., 2010).

Flavonoids are the important plant secondary metabolites, which are necessary for flower coloration, interspecies interaction, disease defense, UV protection and environment challenges (Stefan and Axel, 2005; Chen *et al*., 2014). Flavonoids are synthesized through phenylpropanoid pathway (the partial elements were represented in Fig. 1) and many of its enzymes involved have already been determined. 4-coumarate:coenzyme A ligase (4CL), locating on branch point of the phenylpropanoid derivative biosynthesis, catalyzes the formation of 4-coumarate-CoA from 4-coumarate and coenzyme A (Gross and Zenk, 1974; Lei *et al*., 2011a) and then the 4-coumarate-CoA served as substrates for various important reactions involved in branch metabolism of phenylpropanoid derivative including flavonoids (Dixon and Paiva, 1995; Hahlbrock and Scheel, 1989; Holton and Cornish, 1995). So 4CL is one key enzyme of flavonodids biosynthesis pathway (Fan *et al*., 2007). Many studies revealed that *4cl* gene was a multigene family: Two *4cl* genes are cloned from *Scutellaria baicalensis* Georgi and three *4cl* genes are isolated and characterized in *Hyvrid Poplar* (Allina *et al*., 1998). With further 4CL enzymological identification, genetic mutation and crystal modeling, the studies on the structure and evolution were implemented extensively (Cukovic *et al*., 2001; Schneider *et al*., 2003) and then some highly conserved enzyme active sites residues were revealed, such as Box I (SSGTTGLPKGV) and Box II (GEICIRG) (Stuible and Kombrink, 2001), sbd I (N-terminal domain) and abd II (C-terminal domain) (Ehlting *et al*., 2001).

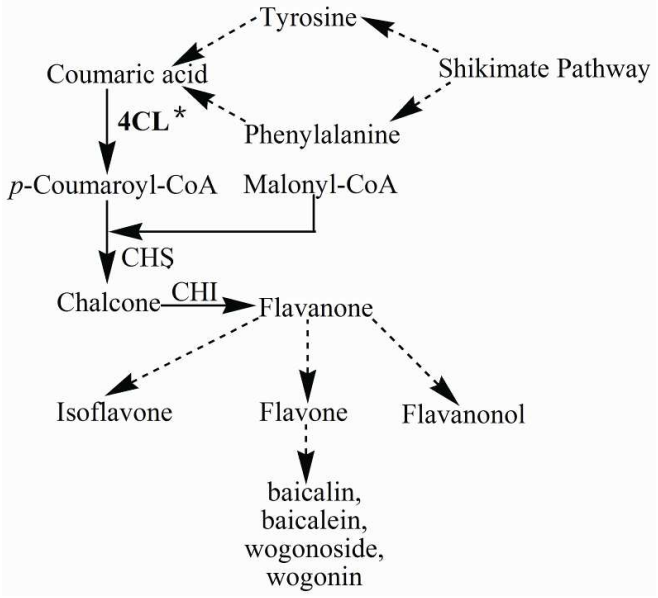

Fig. 1. Flavonoid biosynthetic pathway in *Scutellaria* cells

Individual expression of *4cl* family is regulated by developmental process (Zhao *et al*., 2003), tissue specificity (Kumar and Ellis, 2003) and environmental stress (Ehlting *et al*., 1999), which just answered for the structural diversity of flavonoid compounds and explained their various biological function. Nevertheless, little information is available about molecular structure and physichemical function of 4CL family in *Scutellaria baicalensis* (Lei and Shui, 2014).

S. baicalensis is mainly distributed in East Asia and its dry roots were prevalently used to treat inflammatory and bacterial diseases as old-line China traditional medicine (Yamamoto, 1991; Huang *et al*., 2012; Xue *et al*., 2015). In present study, the bioinformatic analyses of *4cl* family from *S. baicalensis* were completed, which would pave for further studies of physichemical properties of 4CL protein family and its related molecular mechanism of flavonoid biosynthesis.

Materials and Methods

Database Analyses

Two complete sequences with the coding regions (CDS) of *Sb4cl* gene were obtained from NCBI databases: 4CL1 (Accession: AB166767), 4CL2 (Accession: AB166768) and the accession numbers of their corresponding amino acid sequences were BAD90936 (4CL1) and BAD90937 (4CL2).

Bioinformatic Analyses

Comparative bioinformatic analysis of *Sb4cl* was performed at the websites including http://www.expasy.org and http://www.ncbi.nlm.nih.gov. Multiple alignment analysis of the amino acid sequences of Sb4CL and 4CLs from other plant species was finished with Vector NTI Suite 8 (Lei *et al*., 2009). The physicochemical properties was analyzed by ProtParam (Gasteiger *et al*., 2005). The transmembrane helices, subcellular location and hydrophobicity in target proteins were predicted by TMHMM Server v.2.0 (Ikeda *et al*., 2002), TargetP 1.1 Server (Kristin and Siegfried, 2004) and ProtScale (Kyte and Doolittle, 1982) orderly. The motifs of 4CL proteins were searched by ScanProsite. The conserved domains and coiled-coil structures were scanned by CDD (Marchler-Bauer and Bryant, 2004) and COILS (Lei *et al*., 2008) server, respectively. Amino acid sequences of Sb4CL and 4CLs from five species of plants were aligned using ClustalX software (Thompson *et al*., 1997) and subsequently a phylogenetic tree was successfully constructed by Maximum-Likelihood (MP) method with 1000 replicates and another tree was reconstructed by Neighbor-Joining (NJ) with 1000 replicates and meanwhile their reliability of each node was determined by bootstrap calculation using MEGA4.1, respectively (Saito and Nei, 1987; Kumar *et al*., 2008). Finally, the three-dimensional (3D) structures of Sb4CL sequences was modeled based on homological method by Swiss-Modeling (Guex and Peitsch, 1997; Schwede *et al*., 2003; Arnold *et al*., 2006) and then edited and displayed by WebLab ViewerLite 4.2.

Results

Analyses of Structure and Properties

Nucleotide acid sequences of two *4cl* genes were analyzed by the Vector NTI Suite 8 software. They had the same length of Open Reading Frame (ORF), the star codon (ATG) and the stop codon (TGA) and the only

differentiation was that there was one base in the 5' Untranslated Region (UTR) of *4cl2*, but forty-one in *4cl1*. Computed using the online tools ProtParam, some physicochemical parameters were almost identical about 4CL members as shown in the Table 1, such as the formula, isoelectric point (PI), molar extinction coefficient, grand average of hydropathicity (GRAVY) and total number of negatively and positively charged residues and so on.

The tool GOR4 was used for the secondary structure prediction. Sb4CL1 had mixed secondary structure, i.e., random coil, α-helix and extended strand shared a proportion of 47.54, 33.52 and 18.94%, respectively. There was similar composition proportion in Sb4CL2 as shown in Fig. 2 and the coil structures were very high due to abundant hydrophobic praline and flexible glycine amino acids.

Cytological Characterization and Phylogram Analysis

Subcellular localization prediction with the help of online TargetP 1.1 Server inferred that Sb4CL family proteins localized in cytosol without transit peptide. TMHMM Server v2.0 identified no transmembrane region in two 4CL proteins, implying that Sb4CL catalyzed a series of reaction and substrates in cytoplasm without transportation.

After multiple alignments by ClustalX sofeware, two phylogenetic trees of 4CLs were successively constructed from seven plants by MEGA 4.1 with the ME and NJ methods. The most similar result in Fig. 3 showed that *Sb4cl* was most closed relative to each other and the genetic distance was determined to reach 100 nearly.

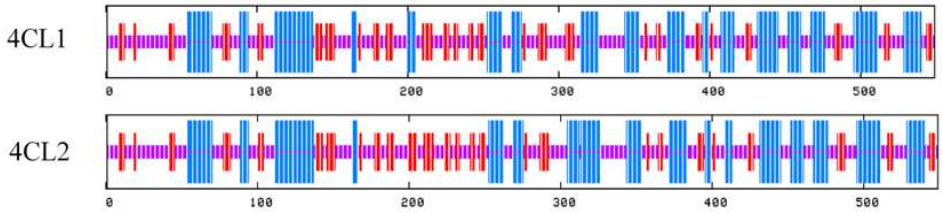

Fig. 2. The secondary structure model of Sb4CL family. The α-helix and extended strand were indicated as and , respectively. Random coil was indicated as

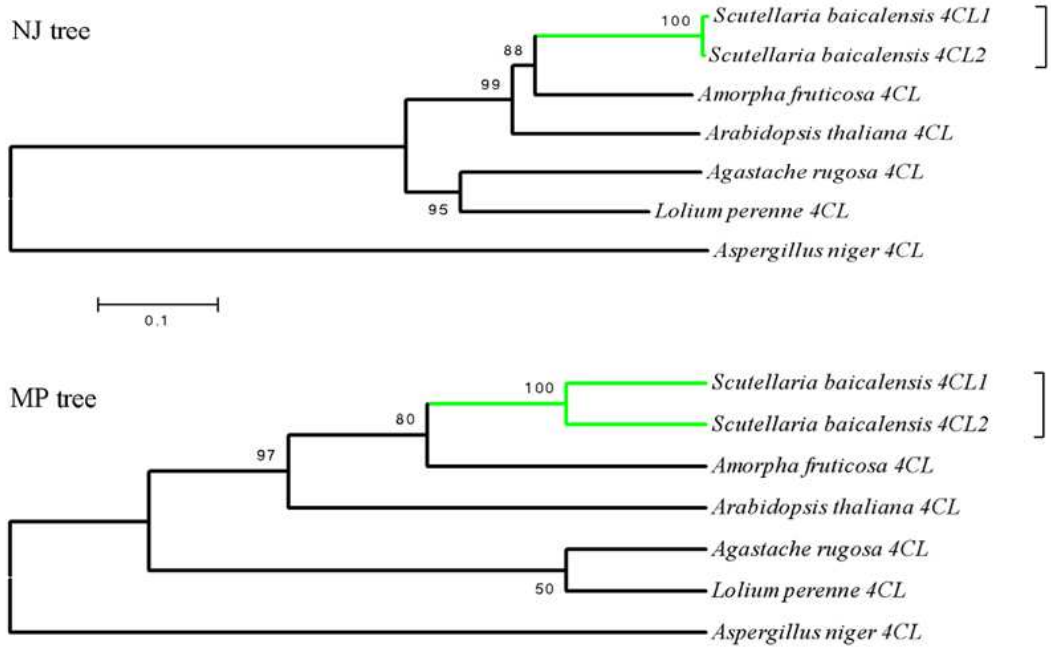

Fig. 3. Molecular phylogram analysis of Sb4CL family and 4CLs from other plants. Phylogenetic trees were constructed by Neighbor-Joining (NJ) and Maximum-Likelihood method, as well as the bootstrap values were showed on branch using MEGA4.1 software. The GenBank accession numbers of the protein sequences used for the phylogenic analysis: *Scutellaria baicalensis* (4CL1: BAD90936; 4CL2: BAD90937), *Agastache rugosa*: AAT02218, *Amorpha fruticosa*: AAL35216, *Lolium perenne*: AAF37732, *Aspergillus niger*: CAK40120, *Arabidopsis thaliana*: Q42524

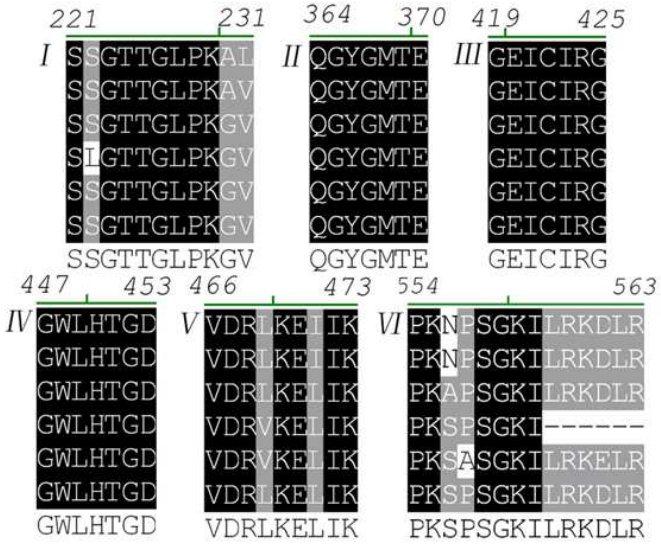

Fig. 4. The multiple alignment of amino acid sequences of Sb4CLs and other plant 4CLs and about six highly conserved regions were shown. The identical sites are shown in white letters and black background; the conservative sites are shown in white letters and gray background; other sites were all shown in black letters and white background

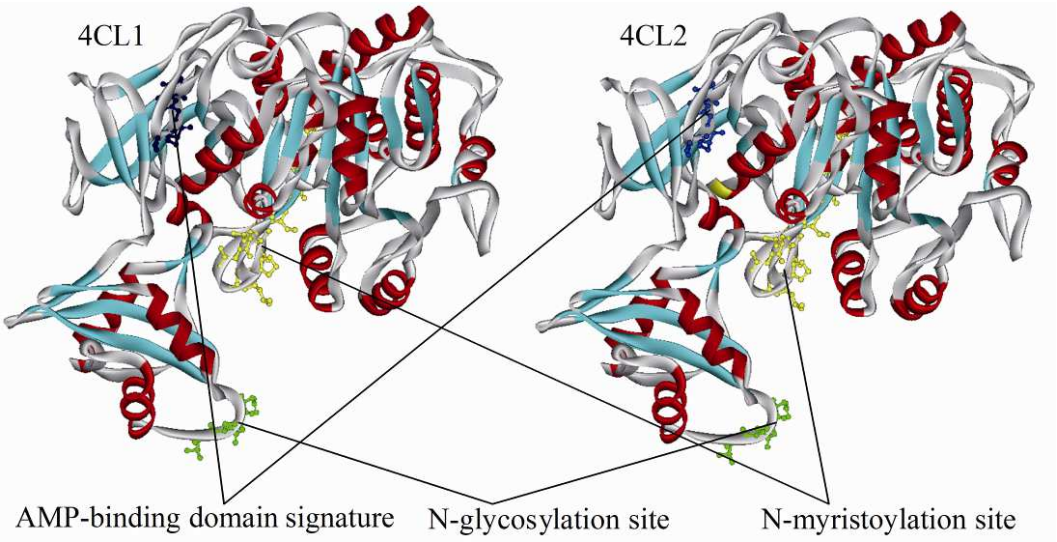

Fig. 5. The 3D structural models of Sb4CL family were established. The α-helixes and β-strands were helix-shaped and wide ribbon-shaped, respectively. Random coils were line-shaped. The three important motifs were marked

Table 1. Analysis of molecular structure and physicochemical properties

Index	4CL1	4CL2
Formula	$C_{2699}H_{4286}N_{694}O_{793}S_{23}$	$C_{2703}H_{4293}N_{697}O_{795}S_{23}$
Molecular weight	59883.2	60012.3
PI	5.35	5.35
Molar extinction coefficient	33975	33975
Estimated half-life	30 hours	30 hours
Instability index	34.55	34.90
Aliphatic index	102.30	101.93
GRAVY	0.109	0.093
Total number of negatively charged residues	67	68
Total number of positively charged residues	50	61

Function Analysis and Three-dimensional Modeling

The tool PROSITE recognized the presence of some motifs with genetic evolutional information and specific biochemical functions, such as an N-glycosylation site (491-494), N-myristoylation site (362-367), AMP-binding domain signature (190-201) in each Sb4CL protein and especially the last two patters were closely related to the important function of 4CL, including modifying myristoyl CoA: Protein N-Myristoyl Transferase (NMT) and acting via an ATP-dependent covalent binding of AMP to their substrate.

The tool CDD recognized the presence of an Acs domain in each Sb4CL protein, suggesting Sb4CL belong to 4CL family. Furthermore, the coiled-coil structure within the Sb4CLs proteins was visualized using COILS online server, polypeptide chain between 368-382aa shaped an obvious coiled-coil structure, confirming there were important function sites located in this region, which was just inlaid within the Acs domain.

Furthermore, the amino acid sequences multi-alignment of Sb4CL family and 4CLs from other four plant species was performed in Vector NTI Suite 8 and Fig. 4 showed the result, in which six highly conserved regions were found orderly from C-terminal to N-terminal: I SSGTTGLPKGV, II QGYGMTE, III GEICIRG, IV GWLHTGD, V VDRLKELIK, VI PKSPSGKILR.

And then, the three-dimensional modeling of the Sb4CLs proteins was visualized using Swiss-Modeling on the basis of the Firefly Luciferase in complex with bromoform and displayed by WebLab ViewerLite. As shown in Fig. 5, some crucial functional domains were marked on the 3-D structure map.

Discussion

Molecular structure and physicochemical properties were analyzed by some bioinformatic tools. Forty-one bases were found in the nucleotide acid sequences of *4cl1* gene, indicating that replication and transcription of *Sb4cl2* gene were impossibly regulated by 5'UTR. Some physicochemical parameters showed high similarity between Sb4CL members and it was important to conclude that *Sb4cl* family was a group of genes with significant genetic conservation and functional association.

The abundant coil structures create effectively links in polypeptide chains and disrupting ordered secondary structure. It appeared that Sb4CL family was associated to ligation of hydroxycinnamate ester and amides. Sb4CL proteins were observed to locate in cytosol, consistent with Geza Hrazdina's report that flavonoid was synthesized in cytoplasmic matrix (Hrazdina, 1992).

4cl gene has been reported in various plants and the researches on its evolutionary were always the hotpoint in the field of the flavonoid metabolic regulation and genetic engineering (Lei *et al*., 2011b). It would be interesting to investigate the *Sb4cl* family evolutionary position in the phylogenetic trees (Huang *et al*., 2008). Belonging to *Scutellaria 4cl* gene family, *Sb4cl1* was most closed relative to *Sb4cl2* in evolutionary level, which also strongly suggested that 4CL was a conserved and committed enzyme of the flavonoid biosynthetic pathway.

Acs domain were identified in each Sb4CL protein, answer for rate-limiting step involved in flavonoids precursor synthesis pathway, i.e., the formation of CoA esters. Additionally, the domain I (i.e., Box I mentioned aboved) was considered as AMP binding motif in 4CL catalytic reaction (Challis *et al*., 2000), which just coincided with the PROSITE prediction that domain I was noted the AMP-binding domain signature. Therefore, domain I SSGTTGLPKGV has become one of the symbols of the adenylate synthase superfamily (Fulda *et al*., 1994; Stuible *et al*., 2000) and meanwhile, domain III (i.e., Box II mentioned aboved) was absolutely conserved in all 4CL proteins, whose central C residue directly participated in catalysis process (Stuible *et al*., 2000).

Conclusion

Based on computational software packages and online servers, bioinformatics analysis can provide useful characterization and prediction of proteins structure and function. In our current study, the nucleotide acid sequences and corresponding amino acid sequences of 4-Coumarate:coenzyme A ligase family from *S. baicalensis* were aligned, analyzed and modeled by some bioinformatic tools and their molecular structures and biochemical functions prediction were obtained as well. The results showed that there was almost no differentiation of molecular structures and physicochemical properties between two members of Sb4CL family, confirming their function relating to flavonoid biosynthesis. The study will be significant in lending theoretical supports for researches of physiochemical properties of 4CL protein and molecular mechanism of flavonoids biosynthesis.

Acknowledgment

This work was supported by National Natural Science Foundation of China (81300035 and 81403044), Collaborative Innovation and Platform Environment Construction Projects of Guangdong Province (2014A030310064 and 2015A050502049), Natural Science Foundation of Guangdong Province (2015A030313520) and Research Project of Traditional Chinese Medicine Bureau of Guangdong Province (20151259).

Author's Contributions

Guoming Li: Performed the study and/or contributed to data analysis and interpretation.

Xiaozhong Lan: Performed the study, wrote the manuscript and/or contributed to data analysis and interpretation.

Xiaorong Shui and Shian Huang: Performed the study and/or wrote the manuscript.

Can Chen: Takes full responsibility for the work as a whole, including the study design, access to data and the decision to submit and publish the manuscript.

Wei Lei: Wrote the manuscript and takes full responsibility for the work as a whole, including the study design, access to data and the decision to submit and publish the manuscript.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

Allina, S.M., A. Pri-Hadash, D.A. Theilmann, B.E. Ellis and C.J. Douglas, 1998. 4-Coumarate: Coenzyme a ligase in *Hybrid poplar*: Properties of native enzymes, cDNA cloning and analysis of recombinant enzymes. Plant Physiol., 116: 743-754. DOI: 10.1104/pp.116.2.743

Arnold, K., L. Bordoli, J. Kopp and T. Schwede, 2006. The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling. Bioinformatics, 22: 195-201. DOI: 10.1093/bioinformatics/bti770

Challis, G.L., J. Ravel and C.A. Townsend, 2000. Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. Biol. Chem., 7: 211-224. DOI: 10.1016/S1074-5521(00)00091-0

Chen, Q., L.P. Pang, S.A. Huang, W. Lei and D.S. Huang, 2014. Effects of emodin and irbesartan on ventricular fibrosis in Goldblatt hypertensive rats. Pharmazie, 69: 374-378. PMID: 24855831

Cukovic, D., J. Ehlting, J.A. VanZiffle and C.J. Douglas, 2001. Structure and evolution of 4-coumarate:coenzyme A ligase (4CL) gene families. Biol. Chem., 382: 645-654. DOI: 10.1515/BC.2001.076

Dixon, R.A. and N.L. Paiva, 1995. Stress-induced phenylpropanoid metabolism. Plant Cell, 7: 1085-1097. DOI: 10.2307/3870059

Ehlting, J., D. Büttner, Q. Wang, C.J. Douglas and I.E. Somssich *et al*., 1999. Three 4-coumarate:coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms. Plant J., 19: 9-20. DOI: 10.1046/j.1365-313X.1999.00491.x

Ehlting, J., J.J.K. Shin and C.J. Douglas, 2001. Identification of 4-coumarate:coenzyme A ligase (4CL) substrate recognition domains. Plant J., 27: 455-465. DOI: 10.1046/j.1365-313X.2001.01122.x

Fan, B.Y., H. Lu and X.G. Jiang, 2007. Review on 4-Coumarate:coenzyme A Ligase (4CL) of vascular plants. Scientia Silvae Sinicae, 43: 96-103.

Fulda, M., E. Heinz and F.P. Wolter, 1994. The *fadD* gene of *Escherichia coli* K12 is located close to rnd at 39.6 min of the chromosomal-map and is a new member of the AMP-binding protein family. Mol. Gen. Genet., 242: 2412-2449. DOI: 10.1007/BF00280412

Gasteiger, E., C. Hoogland, A. Gattiker, S. Duvaud and M.R. Wilkins *et al*., 2005. Protein Identification and Analysis Tools on the ExPASy Server. In: The Proteomics Protocols Handbook, Walker, J.M. (Ed.), Humana Press, Totowa, N.J., ISBN-10: 1588295931, pp: 571-607.

Gross, G.G. and M.H. Zenk, 1974. Isolation and properties of hydroxycinnamate:Co A Ligase from lignifying tissue of *Forsthia*. Eur. J. Biochem., 42: 453-459. DOI: 10.1111/j.1432-1033.1974.tb03359.x

Guex, N. and M.C. Peitsch, 1997. SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. Electrophoresis, 18: 2714-2723. DOI: 10.1002/elps.1150181505

Hahlbrock, K. and D. Scheel, 1989. Physiology and molecular biology of phenylpropanoid metabolism. Ann. Rev. Plant Physiol. Plant Mol. Biol., 40: 347-469. DOI: 10.1146/annurev.pp.40.060189.002023

Holton, T.A. and E.C. Cornish, 1995. Genetics and biochemistry of anthocyanin biosynthesis. Plant Cell, 7: 1071-1083. DOI: 10.2307/3870058

Hrazdina, G., 1992. Compartmentation in Aromatic Metabolism. In: Recent Advances in Phytochemistry, Stafford, H.A. and R.K. Ibrahim (Eds.), Appleton-Century-Crofts, Amsterdam, ISBN-10: 0306442310, pp: 1-23.

Huang, S.A., P.W. Chen, X.R. Shui, Y. He and H.Y. Wang *et al*., 2014. Baicalin attenuates transforming growth factor-β1-induced human pulmonary artery smooth muscle cell proliferation and phenotypic switch by inhibiting hypoxia inducible factor-1α and aryl hydrocarbon receptor expression. J. Pharm. Pharmacol., 66: 1469-1477. DOI: 10.1111/jphp.12273

Huang, S.X., S.L. Hu, X. Sun, Y. Cao and X.Q. Lu *et al*., 2008. Genetic and evolutionary analysis of lignin biosynthase 4CL gene. J. Northwest A&F Univ., 35: 199-206.

Ikeda, M., M. Arai and D.M. Lao, 2002. Transmembrane topology prediction methods: A re-assessment and improvement by a consensus method using a dataset of experimentally-characterized transmembrane topologies. Silico Biol., 2: 19-33. PMID: 11808871

Kristin, E. and H. Siegfried, 2004. InvB is required for type III-dependent secretion of sopA in *Salmonella enterica* serovar typhimurium. J. Bacteriol., 186: 1215-1219. PMID: 14762020

Kumar, A. and B.E. Ellis, 2003. 4-coumarate:coA ligase gene family in Rubus idaeus: cDNA structures, evolution and expression. Plant Mol. Biol., 51: 327-340. PMID: 12602864

Kumar, S., M. Nei, J. Dudley and K. Tamura, 2008. MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief. Bioinform., 9: 299-306. DOI: 10.1093/bib/bbn017

Kyte, J. and R.F. Doolittle, 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol., 157: 105-132. PMID: 7108955

Lei, W. and X.R. Shui, 2014. Study Advance on biosynthesis pathway and biotechnology of baicalin. Northern Horticulture, 22: 185-189.

Lei, W., J.D. Zhang, R.X. Yao and A.M. Qiao, 2011a. Bioinformatic data mining on *ubgat* genes and their encoding proteins in two plants of genus *Scutellaria*. Afr. J. Biotechnol., 10: 4339-4346.

Lei, W., M. Sun, K.M. Luo, X.R. Shui and Y.M. Sun *et al*., 2009. Compute simulations to characterize chalcone synthase from *Scutellaria baicalensis* Georgi. Mol. Biol., 43: 1012-1017.

Lei, W., R.X. Yao, S.H. Tang, X.H. Kang and A.M. Qiao *et al*., 2011b. Isolation and characterization of the anthocyanidin genes *pal*, *f3h* and *dfr* of *Scutellaria viscidula* (Lamiaceae). Genet. Mol. Res., 10: 3385-3402. DOI: 10.4238/2011.November.22.7

Lei, W., S.H. Tang, K.M. Luo and M. Sun, 2010. Molecular cloning and expression profiling of a chalcone synthase gene from hairy root cultures of *Scutellaria viscidula* Bunge. Genet Mol. Biol., 33: 285-291. DOI: 10.1590/S1415-47572010005000031

Lei, W., Tang, S.H., Zhou, Q.G., Shui, X.R., Sun, Y.M., Sun, M., 2008. Bioinformatics analysis of 3-hydroxy-3-methylglutaryl-coa reductase(hmgr) in isoprenoid biosynthesis of mulberry. Acta Sericologica Sinica, 34: 393-399.

Ling, K.H., S.S. Loo, R.R. Mariana, N. Shamsudin and R. Mohamed *et al*., 2007. *In silico* identification and characterization of a putative Phosphatidylinositol 4-Phosphate 5-Kinase (PIP5K) gene in *Eimeria tenella*. Silico Biol., 7: 115-121. MID: 17688436

Marchler-Bauer, A. and S.H. Bryant, 2004. CD-Search: Protein domain annotations on the fly. Nucleic Acids Res., 32: 327-331. DOI: 10.1093/nar/gkh454

Saito, N. and M. Nei, 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol., 4: 406-425. PMID: 3447015

Schneider, K., K. Hövel, K. Witzel, B. Hamberger and D. Schomburg *et al*., 2003. The substrate specificity-determining amino acid code of 4-coumarate:coA ligase. Proc. Natl. Acad. Sci. USA, 100: 8601-8606. DOI: 10.1073/pnas.1430550100

Schwede, T., J. Kopp, N. Guex and M.C. Peitsch, 2003. SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res., 31: 3381-3385. DOI: 10.1093/nar/gkg520

Sivakumar, K., S. Balaji and Gangaradhakrishnan, 2007. *In silico* characterization of antifreeze proteins using computational tools and servers. J. Chem. Sci., 119: 571-579. DOI: 10.1007/s12039-007-0072-y

Stefan, M. and M. Axel, 2005. Flavones and flavone synthases. Phytochemistry, 66: 2399-2407. DOI: 10.1016/j.phytochem.2005.07.013

Stuible, H.P. and E. Kombrink, 2001. Identification of the substrate specificity-conferring Amino Acid Residues of 4-Coumarate:coenzyme A Ligase allows the rational design of mutant enzymes with new catalytic properties. J. Biol. Chem., 276: 26893-26897. DOI: 10.1074/jbc.M100355200

Stuible, H.P., D. Buttner, J. Ehlting, K. Hahlbrock and E. Kombrink, 2000. Mutational analysis of 4-coumarate:coA ligase identifies functionally important amino acids and verifies its close relationship to other adenylate-forming enzymes. FEBS Lett., 467: 117-122. DOI: 10.1016/S0014-5793(00)01133-9

Thompson, J.D., T.J. Gibson, F. Plewniaki, F. Jeanmougin and D.G. Higgins, 1997. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res., 25: 4876- 4882. DOI: 10.1093/nar/25.24.4876

Xue, Y.Q., X.R. Shui, W.Q. Su, Y. He and X.L. Lu *et al*., 2015. Baicalin inhibits inflammation and attenuates myocardial ischaemic injury by aryl hydrocarbon receptor. J. Pharm. Pharmacol., 67: 1756-1764. DOI: 10.1111/jphp.12484

Yamamoto, H., 1991. Biotechnology in Agriculture and Forestry. In: Medicinal and Aromatic Plants III, Bajaj, Y.P.S. (Eds.), Springer Science & Business Media, Berlin, ISBN-10: 364284071X, pp: 502-502.

Zhao, H.Y., J.H. Wei, J. Lu, Y.R. Song and L. Wang *et al*., 2003. cDNA cloning and functional analysis of 4-coumarate — CoA: ligase (4CL) gene in Chinese white aspen. Prog. Nat. Sci., 13: 895-900. DOI: 10.1080/10020070312331344610