Diversity in KCS2 (Ketoacyl-CoA Synthase) of selected plants and its molecular implications: A computational analysis

MONENDRA GROVER 1
National Bureau of Plant Genetic Resources, Pusa Campus, New Delhi 110 012

Received: 16 February 2013; Revised accepted: 30 June 2014

ABSTRACT

The majority of calories in human food are derived from plant fatty acids. Besides, plant fatty acids are also a major component of a variety of products useful to human beings such as paints, cosmetics, biofuels, lubricants, detergents and soaps. Ketoacyl-CoA synthase is a key enzyme involved in the fatty acid elongation in plants. In this study, we have analyzed the diversity in the KCS2 proteins of a selected plant species. We conclude that though there are extensive similarities in the KCS2 proteins studied with respect to total number of negatively charged residues, total number of positively charged residues, and domain organization, there are notable differences for other features such as extinction coefficients, protein stability, kinase specific phosphorylation sites, number of O-GlcNAc sites, predicted sumoylation sites, residues contributing to nuclear export signal and transmembrane helices. These differences may have repercussions for the quantitative efficiency of the 3-Ketoacyl-CoA synthase enzyme which catalyzes the condensation of c2 units to acyl coA during the fatty acid elongation process, and its regulation. This paper showcases molecular implications of diversity in KCS2, which can be used to create a diverse genetic base for engineering KCS2 genes.

Key words: Acyl-CoA, Fatty acids, KCS2, Proteins

MATERIALS AND METHODS

The Arabidopsis KCS2 protein (NM_100303) (AT1G04220) was retrieved from NCBI. This protein was subjected to sequence homology analysis using BLASTP against the nr-Protein database. Some of the hits were selected for further analysis (Table 1).

RESULTS AND DISCUSSION

Fatty acids with chains of twenty or more carbons are termed as VLCFAs (very long chain fatty acids). VLCFAs are precursors of various lipids in plants such as sphingolipids and the phospholipids in the membranes (Schneiter et al. 1996, Devaiah et al. 2006, Dickson et al. 2006), triacylglycerols accumulating in the seeds (Stefansson et al. 1999). In addition to this, the plant fatty acids are also a major component of paints, cosmetics, biofuels, lubricants, detergents and soaps (Ohlrogge 1994). There are two major reasons for mainstream attention to genetic engineering of vegetable oils. First, storage form of fatty acids is a good target for genetic engineering and efforts in this direction are unlikely to have repercussions on the physiology of the plant. Secondly, a vast array of different fatty acid structures with potentially useful properties occur in plants (Voelker and Kinny 2001). Besides, the plant oils can provide viable bio based alternatives to petroleum in the long run.

It is important to understand and analyze genetic diversity which is available to human beings, since modern agricultural practices have led to the erosion of genetic base in the crop plants. In this study, we have analyzed the genetic diversity of 3-Ketoacyl-CoA synthase (encoded by KCS2) from the selected plants in respect to various structural and psycos-chemical parameter and its molecular implications. This will be important in creating a library of KCS2 genes (with known characteristics of the proteins coded by them) which may be used for genetic transformation and thus increase the genetic variability of KCS2 genes.

1 e-mail: monendra_grover@yahoo.com, Present Address: Centre for Agricultural Bioinformatics, Indian Agricultural Statistics Research Institute, Pusa Campus, New Delhi 110 012
August 2014

**Computational Analysis of Diversity of KCS2**

of VLCFAs and their derivatives such as ketones, wax esters, primary alcohols, secondary alcohols, aldehydes and alkanes constitute the cuticular waxes (Post-Beittenmiller 1996, Kunst and Samuels 2003, Samuels et al. 2008). The cuticular waxes ensure protection from UV light (Reicosky and Hanover 1978) and provide safety against lipophilic pathogenic spores and dust (Barthlott and Neinhuis 1997).

The fatty acid elongation in plants proceeds through a series of four reactions: condensation, reduction, dehydration followed by reduction again. In the first condensation step, the C2 carbon moiety is condensed to acyl coA. 3-Ketoacyl-CoA synthase, the 3-Ketoacyl coA is then reduced by the action of 3-ketoacyl-CoA reductase (KCR), which in turn is followed by dehydration of 3-hydroxyacyl-CoA by 3-hydroxyacyl CoA dehydratase and finally there is reduction of trans-2,3-enoyl CoA by trans-2-enoyl CoA reductase. The VLCFAs generated during the above process are then transformed into aldehydes ketones, alkanes and secondary alcohols by decarbonylation pathway or into primary alcohols and wax esters by primary reduction pathway.

(Kunst and Samuels 2003, Samuels et al. 2008). During the fatty acid elongation process, the first step is condensation of C2 units to acyl CoA by KCS. In the Arabidopsis genome a total of 22 KCS genes have been identified (Beisson et al. 2003). When Arabidopsis KCS genes were expressed in yeast, KCS2/DAISY catalyzed the two-carbon elongation of the docosanoic acid (C20 fatty acid) (Franke et al. 2009).

Studies by Lee et al. (2009) indicate that KCS20 and KCS2/...
### Table 6 Phosphorylation sites in the KCS2 proteins

<table>
<thead>
<tr>
<th>Protein id</th>
<th>Plant</th>
<th>Phosphorylation sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>NM_100303</td>
<td><em>Arabidopsis thaliana</em></td>
<td>Ser: 17 Thr: 6 Tyr: 2</td>
</tr>
<tr>
<td>ABV60087.1</td>
<td><em>Gossypium hirsutum</em></td>
<td>Ser: 11 Thr: 6 Tyr: 2</td>
</tr>
<tr>
<td>XP_002283935.1</td>
<td><em>Vitis vinifera</em></td>
<td>Ser: 6 Thr: 6 Tyr: 2</td>
</tr>
<tr>
<td>EAY99126.1</td>
<td><em>Oryza sativa indica</em></td>
<td>Ser: 9 Thr: 7 Tyr: 3</td>
</tr>
<tr>
<td>XP_002437234.1</td>
<td><em>Sorghum bicolor</em></td>
<td>Ser: 9 Thr: 4 Tyr: 4</td>
</tr>
<tr>
<td>ACG36525.1</td>
<td><em>Zea mays</em></td>
<td>Ser: 11 Thr: 2 Tyr: 5</td>
</tr>
<tr>
<td>AAT65207.1</td>
<td><em>Brassica napus</em></td>
<td>Ser: 16 Thr: 8 Tyr: 6</td>
</tr>
</tbody>
</table>

### Table 7 Kinase specific phosphorylation sites in KCS2 proteins

<table>
<thead>
<tr>
<th>Protein id</th>
<th>Plant</th>
<th>Kinases for which sites are present</th>
</tr>
</thead>
<tbody>
<tr>
<td>NM_100303</td>
<td><em>Arabidopsis thaliana</em></td>
<td>PKC, PKA, cdc2, CKI, DNAPK, ATM, CKII, PKG, SRC, p38MAPK, GSK3, cdk5 (12)</td>
</tr>
<tr>
<td>ABV60087.1</td>
<td><em>Gossypium hirsutum</em></td>
<td>CKII, cdc2, PKA, PKC, p38MAPK, DNAPK, RSK, SRC, GSK3 (10)</td>
</tr>
<tr>
<td>EAY99126.1</td>
<td><em>Oryza sativa indica</em></td>
<td>PKC, cdc2, CKII, RSK, p38MAPK, GSK3, cdk5 (12)</td>
</tr>
</tbody>
</table>

DAISY are functionally redundant in the two-carbon elongation to C22 VLCFA that is required for root suberin and cuticular wax biosynthesis. However, their expression is differentially controlled under osmotic stress conditions.

In this study we computationally analyze the properties of the proteins encoded by *KCS 2* gene in Arabidopsis and their homologues in selected plants. Protparam was used to analyze the physico-chemical properties of the KCS2 proteins. It was found that there is slight variation in the number of negatively charged amino acids (48–51).
positively charged residues (62–68) in plants studied. The extinction coefficient of all the proteins varied from 58955 to 68925. The sorghum and maize proteins were found to be unstable in contrast to other proteins studied. The theoretical pl of the proteins studied was found to be broadly similar (Table 3).

The search for domains in the proteins using interproscan revealed that domain organization of all the proteins studied is broadly similar with domains IPR012392, IPR012393, IPR013601, IPR013747, IPR016038, IPR016039 present in all the proteins. The results also show that the proteins are not targeted to mitochondria, chloroplast or peroxisomes. The Signal P analysis showed that none of the proteins are targeted to mitochondria, chloroplast or peroxisomes. The studies on the transmembrane helices show that all the proteins except the B. napus have 10 NESs. The B. napus protein has 2 NESs (Table 10).

O-GlcNAc sites in the proteins were predicted using YinOYang. Predictions for O-GlcNAc attachment sites in the studied proteins are shown in Table 5. The addition of a carbohydrate moiety to the side chain of a residue in a protein chain results in modification of the physico-chemical properties of the protein. Glycosylation is known to result in changes in protein solubility, stability, proteolytic resistance, and local structure (Lis and Sharon 1993, Houssell et al. 1996). The number of potential O-glucosylations is highly variable.

Protein phosphorylation is a ubiquitous mechanism that regulates cellular processes by phosphorylating the serine, threonine or tyrosine residue of a protein. Protein phosphorylation can play an important role in a variety of biological processes, such as, cell cycle progression, transcriptional regulation, signaling transduction and differentiation etc. The number of potential phosphorylation sites was largely similar in the studied plants.

The prediction of Sumoylation sites was done in the selected proteins Small Ubiquitin-like Modifier or SUMO proteins are a family of small proteins that are detached from and attached to other proteins to modify their function. Protein sumoylation can play an important role in a variety of biological processes, such as, cell cycle progression, transcriptional regulation, signaling transduction and differentiation etc. Only the rice and cotton proteins showed the presence of sumoylation sites (Table 8).

The studies on the presence of nuclear export signals show that all the studied proteins have nuclear export signal. The studies on the transmembrane helices show that all these proteins are membrane proteins. All the proteins show the presence of transmembrane helices, with a significant number of amino acids present in the helices (Table 11).

In the present study, we have analyzed the diversity in the KCS2 protein encoded by the KCS2 genes, there are a few similarities in the biochemical and biophysical properties of the proteins encoded by the KCS2 genes, however, there are significant differences. These differences could be related with the fatty acid quantities and quality among plant species and the potential KCS2 genes may be used for genetic engineering of plants for improving oil quality.

REFERENCES
