Evolutionary Profiling of Group II Pyridoxal-Phosphate-Dependent Decarboxylases Suggests Expansion and Functional Diversification of Histidine Decarboxylases in Tomato

Rahul Kumar,* Gitanjali Jiwani, Amit Pareek, Thula SravanKumar, Ashima Khurana, and Arun K. Sharma

Abstract
Pyridoxal phosphate (PLP)-dependent enzymes are one of the most important enzymes involved in plant N metabolism. Here, we explored the evolution of group II PLP-dependent decarboxylases (PLP_deC), including aromatic L-amino acid decarboxylase, glutamate decarboxylase, and histidine decarboxylase in the plant lineage. Gene identification analysis revealed a higher number of genes encoding PLP_deC in higher plants than in lower plants. Expression profiling of PLP_deC orthologs and syntegs in Arabidopsis thaliana (L.) Heynh., pepper (Capsicum annuum L.), and tomato (Solanum lycopersicum L.) pointed toward conserved as well as distinct roles in developmental processes such as fruit maturation and ripening and abiotic stress responses. We further characterized a putative promoter of tomato ripening-associated gene (SlHDC10) operating in a complex regulatory circuit. Our analysis provides a firm basis for further in-depth exploration of the PLP_deC gene family, particularly in the economically important Solanaceae family.

Pyridoxal phosphate (a derivative of vitamin B6) is a pivotal organic cofactor and used by diverse range of enzymes in all living organisms (Facchini et al., 2000). Pyridoxal-phosphate-dependent enzymes catalyze the transfer of amino groups, decarboxylation, and formation of L- and D-amino acids and removal of chemical groups attached to β- and γ-carbon. Their enzymatic versatility arises as a result of their ability of binding of PLP group covalently to the substrate. Their electrophilic catalytic property further assists them stabilize a chemical reaction by stabilizing different carbon ionic intermediates (John, 1995; Schneider et al., 2000).

Depending on distinct structural groups, PLP enzymes are grouped in at least five evolutionary independent families (Milano et al., 2013; Percudani and Peracchi, 2003). One of such important classes is group II PLP-dependent decarboxylase (PLP_deC) enzymes. This class includes aromatic L-amino-acid decarboxylase (AAD), catalyzes the conversion of L-phenylalanine to phenylamine), glutamate decarboxylase (GAD, catalyzes the conversion of L-glutamate to γ-aminobutyric acid (GABA), γ-aminobutyric acid; GABA, γ-aminobutyric acid; GAD, glutamate decarboxylase; GADC, glutamate decarboxylase; HDC, histidine decarboxylase; HMM, hidden Markov models; IAA, indole-3-acetic acid; MS, Murashige and Skoog; PCR, polymerase chain reaction; PLP, pyridoxal phosphate; PLP_deC, pyridoxal-phosphate-dependent decarboxylase; qPCR, quantitative polymerase chain reaction; SDC, serine-like decarboxylase.

Published in The Plant Genome 9
doi: 10.3835/plantgenome2015.07.0057
© Crop Science Society of America
5585 Guilford Rd., Madison, WI 53711 USA
An open-access publication
All rights reserved.

R. Kumar and T. SravanKumar, Repository of Tomato Genomics Resources, Dep. of Plant Sciences, Univ. of Hyderabad, Hyderabad, 500046, India; G. Jiwani, A. Pareek, and A.K. Sharma, Dep. of Plant Molecular Biology, Univ. of Delhi South Campus, New Delhi, 110021, India; A. Khurana, Botany Dep., Zakir Husain Delhi College, Univ. of Delhi, 110002, India; R. Kumar and G. Jiwani contributed equally to this work. Received 28 Mar. 2015. Accepted 31 Aug. 2015. *Corresponding author (rahulpmb@gmail.com).

Abbreviations: AAD, aromatic-L-amino acid decarboxylase; ABA, abscisic acid; BAP, 6-benzylamonopurine; GA, gibberellic acid; GABA, γ-aminobutyric acid; GAD, glutamate decarboxylase; GADC, glutamate decarboxylase; HDC, histidine decarboxylase; HMM, hidden Markov models; IAA, indole-3-acetic acid; MS, Murashige and Skoog; PCR, polymerase chain reaction; PLP, pyridoxal phosphate; PLP_deC, pyridoxal-phosphate-dependent decarboxylase; qPCR, quantitative polymerase chain reaction; SDC, serine-like decarboxylase.
[GABA]), and histidine decarboxylase (HDC, catalyzes the conversion of histidine to histamine) enzymes. Studies on characterization of PLP\_deCs, mainly GADs and AADs, have implicated them in the production of secondary metabolites and flavor volatiles in plants (Akihiro et al., 2008; Clayton, 2006; Facchini et al., 2000; Gallego et al., 1995; Sakai et al., 2007; Sorrequieta et al., 2010; Tieman et al., 2006). Two genes, namely SIGAD2 and SIGAD3, are positively associated with GABA accumulation in maturing fruits of tomato (Akihiro et al., 2008; Gallego et al., 1995). Recently, two serine-like decarboxylases (SDCs) have been implicated in generation of acetaldehydes through decarboxylation and oxidative deamination of phenylalanine, methionine, leucine, and tryptophan (Torrens-Spence et al., 2014). In spite of their importance in plant metabolism, surprisingly, we found limited information available on PLP\_deCs in tomato and other plant species. Moreover, published evidence suggests that there is a scope for further improvement on the existing annotations of previously identified PLP\_deC orthologs, as it was based on a limited sequence data available on plant genomes (Picton et al., 1993; Rontein et al., 2001; Tieman et al., 2006; Torrens-Spence et al., 2014).

In this study, we identified PLP\_deC members from 18 species from plant kingdom, ranging from algae to lower and higher plants. The analysis clearly demonstrated the expansion of class II PLP\_deC family genes during evolution of higher plants. The subgroup HDC showed a clear expansion in tomato. Transcript profiling of these genes identified five ripening-associated HDC genes in tomato. Nonetheless, transcript levels of PLP\_deC genes are also affected by nitrate, phytohormones, and abiotic stresses. Altogether, our results not only improve the current understanding of the mechanisms associated with the evolutionary expansion, sequence conservation, and functional divergence of PLP\_deC genes in tomato but also provide a foundation for further research on them in plants, in general, and Solanaceae family, in particular.

**Experimental Procedure**

**Identification of Pyridoxal-Phosphate-Dependent Decarboxylase**

To identify PLP\_deC genes, protein sequences of full complement of human, *Arabidopsis*, rice (*Oryza sativa* L.), *Escherichia coli*, and yeast pyridoxal-phosphate-dependent decarboxylases (Interpro entry, IPR002129; pfam entry, PF00282) were retrieved from EMBL-EBI webpage (http://www.ebi.ac.uk/interpro/entry/ IPR002129/taxonomy). Additionally, name searches for histidine decarboxylase, aromatic-L-amino decarboxylase, glutamate decarboxylase, and pyridoxal phosphate-dependent decarboxylase were conducted on the Sol Genomics website (http://solgenomics.net/search/loci) for tomato and pepper genes and protein sequences for all the output entries were downloaded. All the unique protein sequences were clubbed with the earlier retrieved proteins and combined sequences were used to generate hidden Markov models (HMM) profile using HMMERv3.1b1 (http://hmmer.janelia.org/) (Supplemental Table S1 in Supplemental File S1). Additionally, protein.fasta files containing all the proteins encoded by the respective genomes, included in this study, were downloaded from either ENSEMBL Plants (http://plants.ensembl.org/index.html) or Phytozome v9.1 (http://www.phytozome.net/). Repository databases. Already generated HMM profiles were used to search the protein.fasta files and the entries listed in output files (score ≥100 and e-value ≤10\(^{-10}\)) were saved. The protein sequences of the listed entries were retrieved by using a custom sequence.extraction perl script. The retrieved protein sequences were then checked for the typical PLP\_deC domain architecture by using pfam (http://pfam.xfam.org/search#tabview = tab1) and SMART (http://smart.embl-heidelberg.de/) tools. The protein sequences lacking the characteristics PLP\_deC domain were discarded and not included in further analyses. Promoter architecture of tomato PLP\_deCs was studied by extracting 1-kb promoter region from their start codon. The cis-elements were analyzed by using online web tools such as PLACE (http://www.dna.affrc.go.jp/PLACE/) and PlantCARE (bioinformatics.psb.ugent.be/webtools/plantcare/html/).

**Multiple Sequence Alignment, Phylogenetic Analysis, and Gene Nomenclature**

ClustalX v2.0 was employed to generate a multiple sequence alignment file of the retrieved PLP\_deC proteins (Thompson et al., 2002). Phylogenetic analysis was performed by using Molecular Evolutionary Genetic Analysis (MEGA 6) tool. Neighbor-joining algorithm with p-distance method and pair wise deletion of gaps, using default parameters was adopted to generate an unreooted phylogenetic tree. A total number of 1000 iterations were tested per analysis and bootstrap statistical analysis for these replicates was employed to test the phylogeny (Tamura et al., 2013). Nomenclature of the identified genes, belonging to the three subcategories, including HDC, AAD, and GAD, was performed on the basis of either their respective positions on the chromosomes from 1 to 12 (first priority) or increasing ENSEMBL PAC IDs used for their annotation. Tomato PLP\_deCs were named based on their position on 12 pseudo molecules, whereas pepper homologs were designated as per their relative homology with tomato genes (Supplemental Table S1 in Supplemental File S1).

**Plant Growth Conditions and Treatments with Phytohormones and Growth Regulators and Abiotic Stresses**

Tomato plants (*S. lycopersicon* 'Pusa Ruby') were grown and maintained under a daily photoperiodic regime of 16 h light and 8 h dark in a greenhouse or culture room, unless mentioned specifically. The temperature of greenhouse was maintained at 28 ± 2°C. Different stages of fruits were scored by tagging them at very young
developmental stage (7 d after pollination with fruit size approximately 1 cm). Fruits were harvested from mature green (MG), breaker (B), 3 d postbreaker (B+3), 7 d postbreaker (B+7), and 20 d postbreaker (B+20) stages, as described previously (Kumar et al., 2011). The harvested tissues were flash frozen in liquid nitrogen to avoid any injury. Three-week-old tomato seedlings grown in culture room were used for rooting host, shoot, and leaf tissues. Various chemical treatment experiments were performed exactly in the same manner as reported earlier (Kumar et al., 2012a, 2015). Ten-day-old seedlings grown in culture room were immersed in the 50-µM solutions of abscisic acid (ABA; Sigma-Aldrich), 6-benzylaminopurine (BAP; Sigma-Aldrich), gibberellic acid (GA; Sigma-Aldrich), and ethylene (ethrel, Sisco Research Laboratories Pvt. Ltd) for two time points, that is, 1 and 3 h. Same stage seedlings treated with water for the same time periods served as their controls. Similarly, 10-d-old seedlings grown on 0.5× Murashige and Skoog (MS) agar were exposed to various stress treatments including cold, desiccation, heat, and salt. For cold stress, culture tubes with tomato seedlings were transferred to a cold chamber maintained at 4 ± 1°C for 2- and 24-h time periods. Similarly, heat stress was given by incubating tomato seedlings in culture tubes in an incubator maintained at 42 ± 1°C, for 30-min and 8-h time points. Seedlings were removed from culture tubes and left on 3 mm blotting paper for 8 h for desiccation stress. For salt stress, seedlings were removed from solid medium without disturbing roots and kept in liquid 0.5× MS medium with 200 mM NaCl for 8 h. Untreated seedlings were used as control for stress treatments. Tissue samples were harvested after the stress treatments and flash frozen in liquid nitrogen. Auxin (50 µM indole-3-acetic acid; IAA) treatment to the KPSC buffer (10 mM potassium phosphate, pH 6.0; 2% sucrose; 50 µM chloramphenicol) treated (12 h) apical portion (10 mm) of 3-d-old etiolated seedlings was performed as described earlier (Kumar et al., 2011). The KPSC buffer was changed every 2 h of the treatment. Seedlings treated with only KPSC buffer (without IAA) for the same time periods served as their controls. Similarly, the available RNA sequencing expression data of tomato (Sato et al., 2012) and pepper (Kim et al., 2014) were also analyzed. The expression profiles of different developmental stages were imported and final heat maps were plotted using gplots package in R. Quantitative polymerase chain reaction (qPCR) analysis was employed to validate the expression of selected tomato PLP, decC genes during different stages of development and also to study their expression under various treatments. In the case of cold and heat stresses, 1 µg RNA each from the two stages of treatments were mixed in equimolar concentration before cDNA preparation. The expression of gene was quantified by SYBR-green-based standard method following the methodology reported earlier (Kumar et al., 2015).

Isolation and Cloning of SHDClO Promoter

Promoter architecture of SHDClO was studied by performing in silico prediction of various cis-regulatory elements present in 1599-bp upstream region, upstream to the start codon. This region was amplified from BAC C08HBA0018C13 using the primers named HDC10-fw: 5′-TTTACGATCCGAGACAAATTTTATCTTGACC-3′ and HDC10-rv: 5′-TTTACGCGGGGAAGTTCTGGCATGGACTAGC-3′ using DNA polymerase enzyme (iMax II; Intron). The amplified PCR fragment was digested with BamHI and Smal restriction enzymes and then purified product was cloned upstream to gus (β-glucuronidase) gene in PB101 vector. Same strategy was used for making two SHDClO deletion:GUS constructs. The sequences of primers used for making these constructs are listed in Supplemental Table S2 in Supplemental File S1. The final constructs were used for subsequent transient as well as stable tomato transformation using Agrobacterium tumefaciens AGL1 strain, as described earlier (Sharma et al., 2009). Briefly, cotyledons from 11-d-old 0.5× MS medium-grown Pusa Ruby seedlings were used for cocultivation with Agrobacterium cultures containing binary vector pBI101-SHDC10::Fl::GUS (Murashige and Skoog 1962). The cocultivated cotyledons were transferred to 1× MS supplemented with 1 mg L⁻¹ transzeatin and 100 mg L⁻¹ kanamycin for regeneration in Petri-plates (9 cm). Plates were cultured under a 16 h light vs. 8 h dark cycle at 28 ± 2°C and explants with callus formation were subcultured onto fresh medium every 15 d. The differentiated shoots were excised from callus and transferred to rooting medium (1× MS supplemented with 100 mg L⁻¹ kanamycin). Those plantlets with healthy shoots and dense roots were transferred to pots containing 50% Soilrite (Kelp- rilite; 1:1:1 ratio of vermiculite, perlite, and Sphagnum moss) mixed with soil for hardening. The young plantlets were transferred to green house (as described earlier) and characterized for the presence of transgene by using nptII-specific primers. Kanamycin-resistant (1× MS supplemented with 500 mg L⁻¹ kanamycin) transgenic lines were propagated for T2 generation and further analysis.

Transient Expression of the SHDClO Promoter Expression Constructs in Nicotiana benthamiana

For transient assays, Nicotiana benthamiana Domin plants were grown in growth chamber at 26 and 22°C day and night temperature cycles in the condition of...
cycle of 16 h light and 8 h dark. For each set of experiment, 6- to 8-wk-grown six-leaf staged plants were used. Mostly, the fourth and older true leaves of tobacco were used for the experiment. Independent *A. tumefaciens* AGL1 cultures containing both full length and deletion promoter constructs were separately grown in YEB medium containing rifampicin and kanamycin at 28°C. The grown culture was then transferred to induction medium (0.5% beef extract, 0.1% yeast extract, 0.5% peptone, 0.1% sucrose, 2 mM MgSO₄, 20 µM acetylsyringone, 10 mM MES pH 5.6) with appropriate antibiotics and grown overnight. Next day, culture were centrifuged and resuspended in infiltration medium (10 mM MgCl₂, 10 mM MES, 200 µM acetylsyringone, pH 5.6, O.D. ≥ 1). This culture was incubated at room temperature with gentle agitation of 20 rpm for 2 h and infused against the lower side of leaf lamina using a syringe with the needle removed. The infiltration was done at 3 to 4 places within a leaf with some distance between each of them. In case of nitrate treatment, the tobacco leaves were detached from the plant after 4 to 5 h to prevent nitrate induction from the soil. Following this, the leaves were given nitrate treatment and the control leaf was incubated in Milli-Q water or KCl solution free from any nitrate contaminations. Ethylene treatment was given to detached tobacco leaves, after agroinfiltration, at different concentrations in air-tight containers for 5 h. Similarly, the control was incubated in Milli-Q water.

**Histochemical and Fluoremetric Analysis of β-glucuronidase Activity**

Histochemical assay for gus gene expression in different organs of transgenic plant was performed as per protocol given by Jefferson and colleagues (Jefferson et al., 1987). The plant organs were dissected and incubated in GUS histochemical buffer [50mM sodium phosphate pH 7.0, 50 mM EDTA pH 8.0, 0.5 mM K₃Fe(CN)₆, and 0.5 mM K₄Fe(CN)₆, 0.1% Triton X-100, 1 mM X-Gluc (5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid), 20% methanol] at 37°C for 16 to 24 h. The chlorophyll was removed from the tissue using a destaining solution of acetone–ethanol (1:3). The activity of GUS enzyme was measured using the same protocol with a few modifications. Frozen tissue (100 mg) was homogenized using a glass rod in a tube with addition of a pinch of glass powder for proper grinding. Next, 150 µL of the extraction buffer (50 mM sodium phosphate buffer pH 7.0, 10 mM disodium EDTA, 0.1% Triton X-100, 0.1% N-lauryl sarcosine, and 10 mM β-mercaptoethanol) was added to the minced tissue, mixed gently and centrifuged at 18,213 g for 20 min at 4°C to pellet the debris. The supernatant was then filtered through 40 filter cell strainer to remove residual debris. The supernatant obtained was quantified for protein by using Bradford’s method (Bradford, 1976). The specific activity of GUS was expressed as nmol of 4-MU (mg protein)⁻¹ h⁻¹ and GUS activity of the control plants was deducted from the experimental samples to obtain the final values.

**Results**

We analyzed the in-house fruit transcriptome microarray data (GSE20720) of five stages during ripening in tomato and identified 145 differentially expressed genes belonging to gene ontology functional category ‘Cellular amino acid metabolic processes’ (Supplemental Fig. S1 in Supplemental File S2). The previously identified histidine decarboxylase (annotated as HDC) gene was present in this category. Chromosome localization study showed that HDC is present on chromosome 8 and flanked by three other highly similar paralogous genes on either side (Table 1). Based on this observation, we further explored the tomato genome and identified the full complement of this gene family in tomato (Fig. 1).

**Identification of Pyridoxal-Phosphate-Dependent Decarboxylase Complement in Plants**

Complementation analysis using reference genome of *Chlamydomonas reinhardtii*, two nonflowering lower plants, *Physcomitrella patens* (Hedw.) Bruch & Schimp. and *Selaginella moellendorffii*, 15 higher plant species, bacteria (*E. coli*), yeast (*Saccharomyces cerevisiae*) and human identified full PLP_deC complements in these species (Fig. 1; Supplemental Table S3 in Supplemental File S1). While both bacterial and yeast genomes encoded only GAD genes, three genes, including one gene each for HDC, GAD, and AAD subclasses were encoded by *C. reinhardtii* genome. The most number of PLP_deC members was identified in tomato (30), whereas the least number (10) was shared by *Arabidopsis* and papaya (*Carica papaya* L.). Surprisingly, 20 homologs were found to be present in *S. moellendorffii* (Supplemental Table S3 in Supplemental File S1). Members of the AAD and GAD subclasses were the major contributors of PLP_deC genes in higher plants. In contrast, HDC subclass was the major contributor of the PLP_deC complement in Leguminaceae and Solanaceae (Fig. 1). A total of 20 HDC genes were found to be encoded by tomato genome, whereas only eight homologs were identified in pepper. Only two HDC genes each could be identified in most of the remaining fruit bearing species. Gene structure analysis of tomato PLP_deCs showed that SIAAD2 and SIAAD3 were devoid of introns, whereas SIAAD5 had the most number of introns (12) (Table 1). Protein sizes of tomato AADs and GADs mostly ranged in between 53 and 57 kDa with an isoelectric point <6.0. In contrast, HDC proteins exhibited great variation in their sizes, ranging from 12 to 53 kDa, and had higher isoelectric points (mostly >6.0) (Table 1).

The syntenic analysis using SynFind (http://genom- evolution.org/CoGe/SynFind) identified seven syntelogs, including two each for AADs (SIAAD3-4: CaAAD3-4) and GADs (SGLGD3: CaGAD3 and SGLGD5: CaGAD4, CaGAD5) and three HDCs (SIHDC1: CaHDC1, SIHDC9: CaHDC9, and SIHDC15: CaHDC15) in pepper and tomato (Supplemental Fig. S2 in Supplemental File S2; Supplemental Table S3 in Supplemental File S1).
Phylogenetic analysis further grouped all PLP_deCs into three subclasses (Fig. 1). The AAD and HDC members of dicot origin largely showed higher sequence similarity with each other than that of monocot origin. In contrast, GAD sequences among higher plants were found to be relatively more conserved (Fig. 1). Rice and Arabidopsis PLP_deCs grouped with their sorghum [Sorghum bicolor (L.) Moench] homologs. Members of lower plants from all three subclasses were well interspersed with their counterparts originated from higher plants (Fig. 1). Furthermore, phylogenetic study among fruit bearing species showed that PLP_deC members belonging to different species of the same taxon such as tomato and pepper from Solanaceae or apple (Malus domestica Borkh.), peach [Prunus persica (L.) Batsch], and strawberry (Fragaria ×ananassa Duchesne ex Rozier) from Rosaceae are comparatively more similar in their protein sequences than their homologs originating from a species from another taxon (Fig. 2). In many cases, more than single homologs for several PLP_deC genes such as peach PpAAD4 with four putative apple homologs (MdAAD1, 5, 6 and 8), pepper CaHDC15 genes with seven putative tomato homologs (SHDC13–19), or papaya CpGAD3 genes with four putative citrus GAD genes (CsGAD4–7) were also identified (Fig. 2).

### Table 1. Characteristics of tomato group II pyridoxal-phosphate-dependent decarboxylases (PLP_deC) and their orthologs and syntelogs in pepper.

<table>
<thead>
<tr>
<th>Sol Genomics ID (version Sl2.40)</th>
<th>Gene name†</th>
<th>Exon/ intron</th>
<th>Protein size (kDa)</th>
<th>Chromosome (Sl2.40)</th>
<th>Cellular localization (TargetP)</th>
<th>cis-elements in promoter‡</th>
<th>Tomato PLP_deC syntelogs in pepper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solyc03g044120.1.1</td>
<td>SlAAD1</td>
<td>2/1</td>
<td>477</td>
<td>53.2 (7.1)</td>
<td>3 Other</td>
<td>ABRE, GA, LECPLEACS2</td>
<td>–</td>
</tr>
<tr>
<td>Solyc04g045020.2.1</td>
<td>SlAAD2</td>
<td>5/4</td>
<td>475</td>
<td>37.4 (6.0)</td>
<td>3 Other</td>
<td>ERE, ABS, LTR</td>
<td>–</td>
</tr>
<tr>
<td>Solyc05g005000.1.1</td>
<td>SlAAD3</td>
<td>1/0</td>
<td>506</td>
<td>56.8 (5.9)</td>
<td>7 Other</td>
<td>ERE, HSE, LECPLEACS, LTR</td>
<td>ABS</td>
</tr>
<tr>
<td>Solyc08g006740.1.1</td>
<td>SlHDC11</td>
<td>7/6</td>
<td>485</td>
<td>55.2 (5.9)</td>
<td>1 Other</td>
<td>Other</td>
<td>–</td>
</tr>
<tr>
<td>Solyc08g006750.1.1</td>
<td>SlHDC2</td>
<td>7/6</td>
<td>503</td>
<td>56.8 (5.1)</td>
<td>3 Other</td>
<td>ABRE, HSE</td>
<td>–</td>
</tr>
<tr>
<td>Solyc08g006760.1.1</td>
<td>SlHDC3</td>
<td>7/6</td>
<td>494</td>
<td>56.2 (5.2)</td>
<td>4 Other</td>
<td>ERE, HSE, LECPLEACS2</td>
<td>–</td>
</tr>
<tr>
<td>Solyc08g006770.1.1</td>
<td>SlHDC4</td>
<td>7/6</td>
<td>498</td>
<td>56.2 (5.7)</td>
<td>5 Other</td>
<td>ARFAT, ABRE, LECPLEACS2</td>
<td>–</td>
</tr>
<tr>
<td>Solyc08g006780.1.1</td>
<td>SlHDC5</td>
<td>7/6</td>
<td>504</td>
<td>57.0 (5.4)</td>
<td>11 Other</td>
<td>ERE, HSE, TGA</td>
<td>CA04g11330 (GaHDC7)</td>
</tr>
<tr>
<td>Solyc08g006790.1.1</td>
<td>SlHDC6</td>
<td>7/6</td>
<td>504</td>
<td>57.0 (5.4)</td>
<td>11 Other</td>
<td>ERE, HSE, TGA</td>
<td>CA04g11330 (GaHDC7)</td>
</tr>
<tr>
<td>Solyc08g006800.1.1</td>
<td>SlHDC7</td>
<td>7/6</td>
<td>500</td>
<td>57.0 (5.4)</td>
<td>11 Other</td>
<td>ERE, HSE, TGA</td>
<td>CA04g11330 (GaHDC7)</td>
</tr>
<tr>
<td>Solyc08g006810.1.1</td>
<td>SlHDC8</td>
<td>7/6</td>
<td>500</td>
<td>57.0 (5.4)</td>
<td>11 Other</td>
<td>ERE, HSE, TGA</td>
<td>CA04g11330 (GaHDC7)</td>
</tr>
<tr>
<td>Solyc08g006820.1.1</td>
<td>SlHDC9</td>
<td>7/6</td>
<td>445</td>
<td>50.2 (6.7)</td>
<td>8† Other</td>
<td>ARFAT, ABRE, HSE</td>
<td>CA08g04940 (GaHDC9)</td>
</tr>
<tr>
<td>Solyc08g006830.1.1</td>
<td>SlHDC10</td>
<td>7/6</td>
<td>445</td>
<td>50.9 (6.8)</td>
<td>8† Other</td>
<td>ARFAT, ABRE, HSE, LECPLEACS2</td>
<td>–</td>
</tr>
<tr>
<td>Solyc08g006840.1.1</td>
<td>SlHDC11</td>
<td>7/6</td>
<td>450</td>
<td>46.1 (6.7)</td>
<td>8† Other</td>
<td>ERE, HSE</td>
<td>–</td>
</tr>
<tr>
<td>Solyc08g006850.1.1</td>
<td>SlHDC12</td>
<td>7/6</td>
<td>450</td>
<td>46.1 (6.7)</td>
<td>8† Other</td>
<td>ERE, HSE</td>
<td>–</td>
</tr>
<tr>
<td>Solyc08g006860.1.1</td>
<td>SlHDC13</td>
<td>5/4</td>
<td>472</td>
<td>53.1 (7.1)</td>
<td>8† Other</td>
<td>GARE, GA, ERE</td>
<td>–</td>
</tr>
<tr>
<td>Solyc08g006870.1.1</td>
<td>SlHDC14</td>
<td>5/4</td>
<td>472</td>
<td>53.1 (6.9)</td>
<td>8† Other</td>
<td>ERE, HSE, LECPLEACS2</td>
<td>–</td>
</tr>
<tr>
<td>Solyc08g006880.1.1</td>
<td>SlHDC15</td>
<td>5/4</td>
<td>450</td>
<td>39.3 (7.3)</td>
<td>8† Other</td>
<td>GARE, GA, ERE</td>
<td>–</td>
</tr>
<tr>
<td>Solyc08g006890.1.1</td>
<td>SlHDC16</td>
<td>5/4</td>
<td>450</td>
<td>39.3 (7.3)</td>
<td>8† Other</td>
<td>ERE, HSE, LECPLEACS2</td>
<td>–</td>
</tr>
<tr>
<td>Solyc08g006900.1.1</td>
<td>SlHDC17</td>
<td>5/4</td>
<td>410</td>
<td>46.2 (6.6)</td>
<td>8† Other</td>
<td>ARFAT, GA, HSE, LTR</td>
<td>–</td>
</tr>
<tr>
<td>Solyc08g006910.1.1</td>
<td>SlHDC18</td>
<td>5/4</td>
<td>456</td>
<td>51.1 (6.5)</td>
<td>8† Other</td>
<td>GARE, LECPLEACS2, ABS</td>
<td>–</td>
</tr>
<tr>
<td>Solyc08g006920.1.1</td>
<td>SlHDC19</td>
<td>5/4</td>
<td>472</td>
<td>53.2 (6.9)</td>
<td>8† Other</td>
<td>GARE, LECPLEACS2, MBS</td>
<td>–</td>
</tr>
<tr>
<td>Solyc12g027900.1.1</td>
<td>SlHDC20</td>
<td>2/1</td>
<td>218</td>
<td>25.4 (10.3)</td>
<td>12 Other</td>
<td>No</td>
<td>–</td>
</tr>
</tbody>
</table>

† Names represented in parentheses depict names used for noted genes at chromosome 8, as noted in column 6, in published literature (Picton et al., 1993; Kumar et al., 2012b; Tieman et al., 2006).

§ ARFAT and TGA, auxin-responsive element; ABRE, abscisic-acid-responsive element; GA and GARE, gibberellin-responsive element; ERE, ethylene-responsive element; LECPLEACS2, ethylene-biosynthesis element; HSE, heat-responsive element; LTR, cold-stress-related element; ABS, drought-responsive element; no, no putative promoter sequence could be retrieved for these genes.

1 and * symbols indicate tandem duplication of two sets of genes at chromosome 8.
Gene Expression Studies on PLP_deC Genes during Plant Development

Previously, a ripening-associated histidine decarboxylase has been described in tomato (Kumar et al., 2012b; Picton et al., 1993). To predict the plausible role of the remaining genes in fruit development and ripening, first we studied transcript profiling of PLP_deC genes using RNA sequencing data of tomato available in public domain. Four HDC genes, including SlHDC9, SlHDC10, SlHDC11, and SlHDC12 showed sharp induction in their transcript accumulation at the early stages of ripening (mostly at breaker stage), whereas SlGAD2 showed such induction even earlier, at mature green stage (Fig. 3). SlAAD5, SlHDC1, SlHDC13, SlHDC14, SlGAD1, and SlGAD5 maintained high expression levels in most of the organs, tissues, and stages studied. SlGAD4 was found
kumar et al.: histidine decarboxylases are expanded in tomato

...transcripts of SlAAD1-3, SlHDC2-5, SlHDC8, and SlHDC20 remained at low level ubiquitously (Fig. 3). Further, expression profiles of seven SlHDC genes (SlHDC6, 9–12, and 16) and two genes each of SlAAD (SlAAD4 and 5) and SlGAD (SlGAD2 and 5) were validated by qPCR (Fig. 4). Despite their ripening-associated nature, SlHDC6 and SlHDC12 exhibited high, whereas SlHDC9 showed moderate transcript levels in at least one of the vegetative tissue (Fig. 4). The qPCR analysis majorly supported the expression profiles observed in RNA sequencing data and established SlHDC10 and SlHDC11 as fruit-ripening-associated genes (Fig. 3, 4).

Three tomato syntelog pairs (SlAAD3:CaAAD3, SlGAD5:CaGAD5, and SlHDC1:CaHDC1a) exhibited almost similar expression patterns, whereas paired members of the remaining four syntelogs (SlGAD1:CaGAD3, SlAAD4:CaAAD4, SlHDC9:CaHDC9,
Figure 3. Expression profiles of (A) tomato, (B) pepper and, (C) Arabidopsis PLP_deC genes during plant development. Heatmaps show expression profiles of members of the three subclasses in various vegetative and reproductive tissues and stages. Color scale represents $\log_2$ (RPKM) values in case of tomato and pepper and $\log_2$ (signal intensities on Affymetrix microarray GeneChip) values in case of Arabidopsis. L, leaf; R, root; S, shoot; F, flower; 1, 2, 3CM, fruit diameter in centimeter; MG, mature green; B, breaker; B.5, and B.10, 5, and 10-d postbreaker; DPA, days postanthesis; COT, cotyledon; HYP, hypocotyl; INF, inflorescence; PL, pollen; SIL, silique.
and SIHDC15:CaHDC15) exhibited mutually distinct expression profiles during fruit development in tomato and pepper (Fig. 4). Whereas transcription of SlGAD1 was found to be inhibited at the ripening initiation in tomato; the expression of CaGAD3 appeared to be induced at the similar stages in pepper. On the contrary, expression of SIHDC9 is ripening linked in tomato while its pepper syntelog CaHDC9 expression remained null at these stages. The other two close homologs, HDC6 and HDC7, showed the highest transcript accumulation in root, whereas their expression was undetectable at fruit ripening stages in both tomato and pepper (Fig. 4). GAD2 homologs showed very high expression levels from mature-green to red-ripe stages in pepper and tomato fruits. Expression profiling of PLP_deC genes in Arabidopsis showed that transcripts of all but AtGAD6 remained at high levels throughout development (Fig. 4). Surprisingly, transcript level of AtGAD6 was highest in pollens, whereas its expression remained low in other tissues and stages (Fig. 4).

Expression Studies on PLP_deC Genes under Hormonal Treatments and Abiotic Stresses

We further preformed qPCR and investigated the effect of different plant hormones and abiotic stresses on the expression of selected tomato PLP_deC genes. While ABA and BAP treatments mostly repressed these genes, GA and IAA treatments, on the other hand, induced their transcript levels (Table 2). In silico analysis of the promoter architecture of tomato PLP_deCs identified many important cis-elements (Table 1). We observed very good correlation between auxin responsive element in
Table 2. Expression profiling of selected group II pyridoxal-phosphate-decarboxylase genes under different phytohormone and stress treatments. Tomato seedlings were subjected to all such treatments following the protocol described earlier (Kumar et al. 2012a). Water-treated seedlings for the same time periods were used as controls.

<table>
<thead>
<tr>
<th>Gene</th>
<th>ABA</th>
<th>BAP</th>
<th>C2H4</th>
<th>GA</th>
<th>IAA</th>
<th>Cold</th>
<th>Desiccation</th>
<th>Salt</th>
<th>Heat</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 h</td>
<td>3 h</td>
<td>1 h</td>
<td>3 h</td>
<td>1 h</td>
<td>3 h</td>
<td>1 h</td>
<td>3 h</td>
<td>1 h</td>
</tr>
<tr>
<td>SIAAD4</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>SIAAD5</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>SIGAD5</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>SIGAD9</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>SIGAD6</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>SIGAD10</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>SIGAD11</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>SIGAD12</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
</tbody>
</table>
| SIVAAD4, SIVAAD5, SIGAD5, SIHDC9, SIHDC10, SIHDC12, and SIHDC16 promoters with their induction by this hormone (Table 1, 2). Despite the presence of ethylene-responsive elements in several tomato PLP-deC promoters, ethylene exerted a limited effect on the modulation in their transcript accumulation, except SIHDC10 and SIHDC11, which were elevated by its treatment (Table 1, 2). Transcript level of SIHDC10 was found to be elevated under most treatments except under ABA, desiccation, and salt stresses. Desiccation predominantly had a negative effect, similar to ABA treatment, on the expression of most of genes studied (Table 2). Overall, stress treatments downregulated the expression of PLP-deC genes in tomato. However, a few exceptions, such as induced expression of SIGAD2, SIHDC10, SIHDC11, and SIHDC16 under cold and SIGAD2, SIHDC6, and SIHDC9 under salt stress, were also observed (Table 2). Despite the presence of heat shock element in the promoters of most of tomato PLP-deCs, SIAAD5, SIHDC1, SIHDC12, and SIHDC16 were the only genes that showed a change (mild repression) in their transcripts after heat stress (Table 1, 2).

On the basis of previous studies (Turano and Fang, 1998; Wang et al., 2000), we selected three different concentrations (250 µM, 5 mM, and 10 mM) of potassium nitrate (KNO₃) and analyzed their effect on the transcript accumulation of these genes and found that such nitrate treatment induces several tomato PLP-deC genes, except SIGAD2 and SIGAD5 (Fig. 5). Transcript levels of SIAAD1, SIAAD6, SIHDC1, SIHDC9–12, and SIHDC16 were found to be induced in all nitrate treatments. We noticed an increase in transcript level of SIHDC6 only in 250 µM nitrate treatment, whereas higher nitrate level (10 mM) was found to inhibit its expression (Fig. 5).

Functional Characterization of SIHDC10 Promoter

We next characterized putative promoter of SIHDC10 by retrieving 1.6-kb upstream sequence. Only 1.6-kb region was selected, as information of further upstream sequence was not available at the time of cloning. Besides many other cis-elements, in silico examination of this sequence showed presence of one each of auxin-responsive element (AuxRE, at −1282 bp), putative fruit-specific element (TCCAAAA at −444 bp), and putative nitrate response element (TGACCTTT, at −909 bp), two ethylene biosynthesis elements LECPLEACS2 (at −1480 and −656 bp), three each MADS-box transcription factors binding CARG elements (at −1587, −994, and −718 bp) and ethylene-responsive element (at −402, −390, and −211 bp). In contrast to the very high levels of SIHDC10 transcripts, surprisingly, we could not detect an equivalent GUS activity in fruits of SIHDC10–Fl::GUS transgenic fusion transgenic tomato lines. GUS accumulation was found to be higher in mature green fruits than red fruits. Additionally, roots of these transgenic plants were found to have maximum GUS expression (Supplemental Fig. S3A–C in Supplemental File S2). To find out the reason for these unexpected results, we further made two deletion constructs to identify the minimum promoter sequence (Fig. 6A; Supplemental Fig. S3A in Supplemental File S2). SIHDC10–DII::GUS and SIHDC10–DII::GUS deletion constructs in tobacco leaves resulted in an enhanced activity of shorter promoter region under both nitrate and ethylene treatments. At least three times more GUS accumulation was observed in SIHDC10–DII::GUS (−968 to −1599 bp deletion) or SIHDC10–DII::GUS (−668 to −1599 bp deletion) infiltrated tobacco leaves in comparison to those infiltrated with SIHDC10–Fl::GUS construct. Similarly, the enhanced promoter activity, at least five times and
10-fold higher, was recorded in SlHDC10-DI::GUS and SlHDC10-DII::GUS infiltrated leaves, respectively, under ethylene treatment (Fig. 6A, 6B).

**Discussion**

Type II PLP\_deC enzymes, especially belonging to AAD and GAD subclasses, play important roles in amino acid metabolism during plant growth and development (Akihiro et al., 2008; Facchini et al., 2000; Gutensohn et al., 2011; Tieman et al., 2006). These enzymes are known to have stringent substrate specificities, which suggests toward their functional evolution from a common evolutionary origin. The availability of limited information on their protein sequences and biochemical properties has prevented an understanding their actual roles of many PLP\_deCs in plant development, to date. In the present study, we identified full complement of PLP\_deC genes in many plants (both lower and higher plants) belonging to different taxa and classified them in GAD, AAD, and HDC subclasses (Fig. 1, 2; Table 1). The nomenclature of these genes is strictly based on their high sequence homology, however, we cannot rule out the possibility that some of these genes might have evolved for diverse functions. For example, the already characterized LeAADC1A and LeAADC2, annotated as SIHDC19 and SIHDC6, respectively, in the present study, do not act on histidine but prefer tyrosine as their substrate (Tieman et al., 2006). Further, these tomato enzymes are reported to share higher homology (57%) to the characterized plant AtSDCs than AADs (10–15% identity). Similarly, many SDCs, initially annotated as HDCs, upon biochemical analysis, were found to prefer serine rather than histidine (Rontein et al., 2001). Recent biochemical characterization of SDC-like enzymes from *Medicago truncatula* Gaertn. and *Cicer arietinum* L. further indicates toward their novel aldehyde synthase activity (Torrens-Spence et al., 2014). Therefore, despite of their high sequence similarity, PLP\_deCs have evolved for diverse substrate specificities and their precise function will be known only after in-depth biochemical characterization (Torrens-Spence et al., 2014).

Gene identification analysis showed presence of maximum 30 PLP\_deC members in tomato, whereas papaya and *Arabidopsis* genomes were found to encode the minimum set of such genes among 15 plant species used in this study (Fig. 1). Presence of more PLP\_deC genes in all plant genomes than in bacteria, yeast, and *Chlamydomonas* is in accordance with the earlier results and supports that number of these genes increases with the increased complexity of genomes (Facchini et al., 2000). However, a few exceptions for the ratio of genome size and number of encoded group PLP\_deC genes exist in organisms such as *Chlamydomonas*, pepper, and *Selaginella* (Supplemental Table S3 in Supplemental File S1). This presumably supports the earlier observation that the number of such genes in a genome also depends on the nutrient requirements of organisms and the external environment might have contributed to their evolution and diverse substrate preferences (Facchini et al., 2000).

The comprehensive phylogenetic analyses of PLP\_deC genes are further in accordance with the divergence time of different groups during evolution of plants, as land plants are expected to have diverged from green algae 750 million yr ago, whereas lower land plants are suggested to have diverged from their flowering counterparts by more than 400 million yr on evolutionary time scale (Fig. 1, 2) (Guo et al., 2013; Nickrent et al., 2000). Comparatively lower degree of sequence similarity of PLP\_deC members...
between monocots and eudicots, than that between members of either eudicots or monocots, further supports the earlier observations on their divergence as these two groups are proposed to have diverged at least 150 million yr ago (Fig. 1) (Chaw et al., 2004; Guo et al., 2013; Salinas et al., 2012). Further, predicted protein size of PLP_deCs is in accordance with the previously published reports (Molina-Rueda et al., 2010; Tieman et al., 2006).

Presence of seven syntelog pairs (SlAAD3-4:CaAAD3-4; SlGAD3:CaGAD3, SlGAD5:CaGAD4, CaGAD5, SIHDC1:CaHDC1, SIHDC9:CaHDC9, and SIHDC15:CaHDC15) in tomato and pepper (Fig. 2) suggests that these genes might have originated from a common ancestor before divergence of Solanum and Capsicum lineages (Wu and Tanksley, 2010). Two pepper syntelogs (CaGAD4 and CaGAD5) for SlGAD5 and presence of seven putative tomato homologs (SIHDC13-19) for CaHDC15 suggests that the additional homologs in either species would have evolved through gene duplication event after the speciation (Kim et al., 2014). Altogether, gene identification and phylogenetic analyses revealed that HDC subclass has moderately expanded in most of Leguminaceae and Solanaceae members; however, in tomato, mainly tandem duplication events have contributed in their abrupt expansion (Fig. 1, 2).

The expression data of tomato syntelogs in pepper is very interesting as the conservation and divergence of their function may contribute to quantitative and qualitative differences observed under climacteric and nonclimacteric ripening programs. The similar expression pattern for SlGAD1:CaGAD3, SIHDC9:CaHDC9, and SIHDC15:CaHDC15 syntelogs during fruit developmental stages in both tomato and pepper suggests conservation in their functions in the two species (Fig. 4). In contrast, nonsimilar expression profiles of SlGAD1:CaGAD3, SlAAD4:CaAAD4, SIHDC9:CaHDC9,
and SIHDC15:CaHDC15 syntelogs indicate that these genes have evolved to perform different functions in the two species (Kim et al., 2014). Lack of any syntelog for fruit-specific ripening-associated SIHDC10 and SIHDC11 and almost undetected levels of CaHDC9 transcripts in pepper is interesting and indicates that function of these genes is not required for maintaining fruit quality in the latter species (Fig. 4).

Presence of at least four ripening-induced genes (SIHDC9–12) (Fig. 3, 4) further indicates that these genes might have overlapping functions during fruit ripening in tomato (Tiemann et al., 2006). Histamine levels are known to increase during fruit ripening (Martin et al., 2005; Wojtaszek, 2003). Further, presence of more HDC genes in Solanaceae could be a reason for the relatively higher amount of histamine (mostly >1.5 mg 100 mg⁻¹ fresh weight) reported in these plants (Feldman, 1983; Kumar et al., 2009). Presence of fruit-specific elements viz. TCTTCACA and TCCAAAA in the putative promoters of SIHDC10 and SIHDC11, respectively, supported the expression data and provided an additional evidence for their fruit-specific nature (Table 1) (Yin et al., 2009; Zhao et al., 2009). The higher GUS expression in mature green, whereas very low in red fruits in SIHDC10-Fli:GUS tomato transgenic lines, is intriguing and could be due to either higher nitrate levels in mature green tomatoes or an insufficient length of the putative promoter used, respectively, and most likely be corrected by using a >1.6-kb promoter sequence (Schaart et al., 2002). Moreover, ectopic expression of GUS in the roots could be due to exposure of roots to nitrate source present in soil (Wang et al., 2000). Previous evidence in tomato, tobacco, and Arabidopsis further support the present findings (Eyal et al., 1993; Twell et al., 1990). The analysis of the GUS activity of two deletion constructs established that nitrate responsiveness of SIHDC10-DII:GUS deletion construct could be due to some yet to be characterized nitrate-response element present in the first 668-bp region of this promoter (Das et al., 2007).

The results presented in this study further established that members of PLP_deC family are under the control of multiple signals as transcript level of several members of this gene family in tomato were found to be affected by plant hormones such as auxin, ABA, ethylene, cytokinin, and gibberellins and abiotic stresses including cold, heat, salt, and desiccation and supports earlier findings (Gutensohn et al., 2011; Hyun et al., 2014; Kanwal et al., 2014; Kumar et al., 2012a, 2014; Picton et al., 1993). High correlation between auxin-induced increase in transcript levels of these genes and AuxRE (ARFAT) in their promoters is in accordance with previously published studies (Table 1, 2; Guilfoyle and Hagen, 2007; Ulmasov et al., 1997). However, lack of same level of tight correlation between presence of other phytohormone-responsive elements, vis-à-vis expression of genes under specific phytohormone treatment, suggest toward involvement of complex molecular mechanisms underlying their transcriptional regulation.

Overall, the present study provides an initial glimpse into the evolution of PLP_deC gene family members in plants and improves the current understanding on their plausible role during plant growth and development. Moreover, the expression data under various phytohormones, nitrate treatments, and abiotic stresses in tomato suggests their involvement in contributing toward modified metabolic activities under such conditions in plants. In silico promoter data and functional characterization of SIHDC10 promoter further supports the expression data and suggests that these genes operate under complex molecular circuits. Finally, the comprehensive identification, expression, and functional characterization of group II PLP_deC genes in tomato will further help in assigning them functions in Solanaceae, in general, and tomato, in particular.

Acknowledgments
This work was financially supported by grants received from the Department of Science and Technology (DST), and Department of Biotechnology (DBT), Government of India. GJ, AP, TSK acknowledge CSIR, UGC, and DST, respectively, for the fellowship granted during their tenure as research fellows. RK is grateful to DST for INSPIRE-Faculty Award (grant number IFA-LSPA-15). RK also thanks Prof. R. P. Sharma, Department of Plant Sciences, University of Hyderabad, for his support in carrying out the research work. Authors are grateful to Dr. Hiriyadesh Prakash, University of Hyderabad, and Dr. Pavel Kerchev, Department of Plant System Biology/VIB, University of Gent for critical reading of the manuscript. Authors acknowledge the Solanaceae Genomics Network for the use of Tomato BAC Sequence Database to retrieve putative promoter sequences.

References


