Plant Science 254 (2017) 70–81

Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

Dynein light chain family genes in 15 plant species: Identification, evolution and expression profiles

Jun Cao[a,*,1], Xiangyang Li[b,1], Yueqing Lv[a]

[a] Institute of Life Sciences, Jiangsu University, Xuefu Road 301, Zhenjiang 212013, Jiangsu, PR China
[b] Industrial Crop Institute, Henan Academy of Agricultural Sciences, Zhengzhou 450002, Henan, PR China

ARTICLE INFO

Article history:
Received 14 July 2016
Received in revised form 2 October 2016
Accepted 31 October 2016
Available online 9 November 2016

Keywords:
Dynein light chain
Evolution
Expression
Network analysis
Plant

ABSTRACT

Dynein light chain (DLC) is one important component of the dynein complexes, which have been proved involving in a variety of cellular functions. However, higher plants lack all other components of the complexes except DLCs, suggesting that in plants, the DLC protein does not carry out the same function as it in animals. Therefore, the function of this family in plants is mysterious. In this study, we investigated the *DLC* gene family in 15 plant species and analyzed their expression profiles. In total, 128 *DLC* genes were identified from the 15 studied plant species and were divided into eight groups by their phylogenetic relation. Highly conserved gene structure and motif arrangement was discovered within each group, indicating their functional correlation. Genetic variation and recombination events were also detected in *DLC* genes. Through selection analyses, we also identified some significant site-specific constraints in most of the *DLC* paralogs. In addition, *DLC* genes presented various expression profiles in different development stages, or under different abiotic stresses or phytohormone treatments. This may be associated with a variety of *cis*-elements responding to stress and phytohormone in the upstream sequences of the *DLC* genes. Functional network analysis exhibited 123 physical or functional interactions. The results provide a foundation for exploring the characterization of the *DLC* genes in plants and offer insights for additional functional studies.

1. Introduction

Dyneins are large multi-subunit motor complexes, which contain two dynein heavy chains (DHCs), some dynein intermediate chains (DICs), several light intermediate chains (LICs), and a group of dynein light chains (DLCs) [1]. DHC can be classed either as an axonemal dynein, involved in the movement of flagella and cilia, or cytoplasmic dynein. The cytoplasmic dynein can further subdivided into cytoplasmic dynein 1, involved in intracellular movements, and cytoplasmic dynein 2 which functions in intraflagellar transport. Dimerisation of the DHCs is aided by various DICs, LICs and DLCs [2]. DLC was first identified as a subunit of *Chlamydomonas* flagellar dynein [3,4]. Subsequent studies demonstrated that it is also present in the cytoplasmic organisms or tissues that lack flagella and cilia in various kinds of eukaryotes [5]. The DLC is a subunit with about 8–30 kDa. That is highly conserved in eukaryotic genomes [6]. In addition, two copies of DLC exist in each dynein complex, and they can be bound to the motor via DICs.

DLC proteins usually function as a dimer in a variety of cellular processes. This dimer can be converted into monomers by phosphorylation of Ser-88 residue or low pH [7–9]. Several studies have shown that the DLC can interact with tubulin and regulate microtubule dynamics [10,11]. DLC phosphorylation of S82 residue stabilizes microtubule and then regulates mitochondria from hypoxia damage [12]. In *Drosophila* and *C. elegans*, deletion or inhibition of the *DLC* gene usually induces apoptosis, cell death and early embryonic lethality [13–15]. Deletion of the *DLC* gene in *Yarrowia lipolytica* cell can result in mis-localization and abnormal peroxisome morphology, implying that DLC is involved in peroxisome biogenesis and function [16]. In neuron development and mitosis, DLCs also play key roles [17–20]. Moreover, because of the feature of binding viral proteins, such as, human foamy retrovirus Gag protein [21], rabies P proteins [22,23], p54 protein of African swine fever virus [24] etc., DLC proteins can also be used as a gene delivery vector into the nucleus [25].

Besides the DLC sequences, flowering plants lack all other components of the dynein motor. It suggests that plant DLC protein might perform divergent functions as it does in animals [2]. There are six *DLC* genes in *Arabidopsis*. The only identified function is interacting with Sirevirus Gag protein [26]. No additional cellular functions have been proposed for this family. The complete genome

* Corresponding author.
E-mail address: cjinfor@163.com (J. Cao).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.plantsci.2016.10.011
0168-9452/

sequencing of some models plant species provides the basis for comparative genomic studies for the *DLC* gene family. Considering their significant developmental roles, it is of considerable interest for us to investigate the evolution of plant *DLC* genes. In this study, we identified the *DLC* gene family from 15 plant species, with the range of one species comprises one to 24 genes. Besides, we also performed integrated analyses to unravel the evolutionary mechanisms of the plant DLC family. It provided a useful basis for further functional studies of this gene family in the plant kingdom.

2. Materials and methods

2.1. DLC sequence retrieval and identification in 15 plant species

To identify potential *DLC* genes in 15 completely sequenced plant genomes, we first used the six *Arabidopsis* DLC protein sequences [26] as queries to perform BLAST searches in the phytozome database (http://www.phytozome.net). The Pfam database [27] was also used to confirm the encoded DLC. Afterward, the ProtParam tool (http://web.expasy.org/protparam) and the CELLO v2.5 server (http://cello.life.nctu.edu.tw) [28] were used to predict the physicochemical parameters and subcellular localization of the DLC proteins, respectively.

2.2. Phylogenetic analyses of the DLC gene family

To further explore the evolutionary relationship of the plant DLCs, multiple sequence alignment was carried out on the predicted amino acid sequences by the muscle method [29] with default settings. Phylogenetic analyses of the output alignment sequences were performed with neighbor-joining (NJ) and maximum likelihood (ML) methods using MEGA ver.5 [30]. The NJ tree was constructed with 5000 bootstrap replications. The treatment for gaps data is pairwise deletion and the substitution model was *p*-distance. For the ML tree, distance matrices were based on the Jones-Taylor-Thornton (JTT) substitution model. Two thousands bootstrap replicates were carried out to calculate the relative support for branches of the inferred phylogenetic tree.

2.3. Exon-intron organization and motif composition analysis

DLC gene organization was analyzed by comparing their coding and genomic sequence information in the Phytozome (http://www.phytozome.net). In addition, MEME program (http://meme.sdsc.edu) [31] was used to identify finer motif composition in the candidate plant DLC proteins. The maximum number of motifs is 10.

2.4. Genetic variation and recombination analysis

Nucleotide diversity was estimated with DnaSP V5.0 [32]. Next, K-Estimator 6.0 [33] was used to calculate the synonymous (*Ks*) and non-synonymous substitution rates (*Ka*). Potential recombination events between divergent nucleotide sequences were explored by the program RDP ver.3.44 [34], which embeds different methods for detecting recombination signals. In this study, RDP [35], Geneconv [36], and MaxChi [37] methods were used to detect these signals. The highest acceptable P value cutoff was placed at 0.05. Significance was evaluated with 100 permutation tests.

2.5. Site-specific selection assessment and testing

K_a/K_s values are used to estimate two types of substitutions events by calculating the synonymous rate (K_s) and the non-synonymous rate (K_a), at each codon site. To identify which residue site of DLC proteins in different group is evolutionally selected, the Selecton Server (http://selecton.tau.ac.il/) [38] was used to calculate their K_a/K_s values. This server implements several evolutionary models that describe in probabilistic terms how the characters evolve. In this study, three of the evolutionary models [M8 ($\omega_s \geq 1$), M7 (beta) and M5 (gamma)] were used. These models use different biological assumptions to test the data and to assume a statistical distribution for heterogeneous K_a/K_s values among sites. The distributions are approximated using eight discrete categories and the K_a/K_s values are achieved by calculating the expectation of the posterior distribution [38]. Three-dimensional structure of a DLC protein (Glyma13G277800.1) on Group V was predicted from the Phyre2 Server (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) [39]. Effects of point mutations on folding stability were estimated with I-Mutant2.0 [40]. Free energy change value [$\Delta\Delta$G (kcal/mol)] was used to infer the protein stability. When $\Delta\Delta$G is less than zero, the stability of the mutant protein is reduced; on the contrary, when $\Delta\Delta$G is greater than zero, the stability is increased.

2.6. Microarray-based expression analysis

Microarray data reported by Laubinger et al. [41] and Jain et al. [42] were used for the expression analysis of the *Arabidopsis* and rice *DLC* genes in different development stages, respectively. Expression data were normalized and viewed in the Genesis (v 1.7.6) program [43].

2.7. Plant treatment, RNA isolation, quantitative real-time PCR (qRT-PCR)

Ten-day-old *Arabidopsis* and rice seedlings were used to test the expression profiles of *DLC* genes under cold, drought, IAA, ABA, and salt stresses. The seedlings were stored at 4 °C for 3 h and dried in a folded tissue paper at 23 ± 1 °C for 3 h for cold and drought conditions, respectively. 10 μM IAA and 150 mM NaCl were used to treat *Arabidopsis* and rice seedlings for 24 h. For *Arabidopsis* heat treatment, the seedlings were kept in 42 °C for 3 h. Each sample was tested in triplicate biological replicates. Control (CK) seedlings were normally grown at 23 ± 1 °C with a photoperiod of 14 h light and 10 h dark. Total RNA was extracted with the TRIzol® total RNA extraction kit (Sangon). RNase free DNase-I was used to remove genomic DNA. Next, we used M-MLV (TakaRa) to perform reverse transcription, followed by quantitative assays of each diluted cDNA using an ABI 7500 sequence detection system. All *Arabidopsis* and rice *DLC* genes were performed for qRT-PCR analysis, and their primers are listed in Table S1. *Actin* genes (*AT1G49240* for *Arabidopsis* and *LOC_Os03g61970* for rice) were used as the endogenous control. Finally, we used $2^{-\Delta\Delta CT}$ method [44] to calculate the relative expression level of *DLC* genes. *t*-Test was used to perform a significant analysis.

2.8. Cis-regulation element analyses of the DLC promoters

To define the transcription start site (TSS) of each *DLC* gene in *Arabidopsis* and rice, we first collected their expressed sequence tag information. TSS positions of *DLC* genes were used as references to determine their upstream promoter sequences. In this study, 1000-bp upstream promoter sequences were acquired for further analyses. In addition, some abiotic stress- and phytohormone-responsive elements were identified in the *Arabidopsis* and rice *DLC* promoter regions using PLACE (http://www.dna.affrc.go.jp/PLACE/signalscan.html) [45].