Plant Gene 14 (2018) 55–63

Contents lists available at ScienceDirect

Plant Gene

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

# Diversification of *SQUAMOSA* promoter binding protein-like (SPL) genes by changes of miR156/529 binding sites in land plants

Shu-Dong Zhang[a,b], Li-Zhen Ling[a,c,*]

[a] Department of Life Science, Liupanshui Normal University, Liupanshui, Guizhou 553004, China
[b] Germplasm Bank of Wild Species, Kunming Institute of Botany of the Chinese Academy of Sciences, Kunming 650201, China
[c] BGI-Yunnan, BGI-Shenzhen, Kunming 650106, China

ARTICLE INFO

*Keywords:*
Functional diversity
Gene duplication
Land plant
SPL gene family
Sequence divergence
Structural divergence
miR156/529

ABSTRACT

Gene duplication can increase functional diversity of a gene family. SQUAMOSA promoter binding protein-like (SPL) genes encode plant-specific transcription factors. The SPL genes are regulated by *miR156* and *miR529* microRNA families. However, little is known about how and if duplicated SPL genes (paralogs) have retained active miR156/529 binding sites. Here, we analyzed 136 SPL genes from eight representative land plants, in which 70 of them originated from several large-scale duplication events, 40 ones constituted 32 duplicated target pairs. In this study, these target pairs were divided into divergent pattern (D-type) and same pattern (S-type). The D-pattern included loss/gain of miRNA binding sites on one paralog (D-1) and the mutation on miRNA binding sequences (D-2). Further study indicated that D-type gene pairs had the higher sequence divergence and evolutionary rate than S-type ones. Meanwhile, D-type gene pairs exhibited the different expression patterns, whereas S-type gene pairs had the similar expression patterns. These results revealed that the different divergence of miR156/529 binding sites after gene duplication might contribute to increasing functional diversity of SPL genes. In addition, gene structural analysis suggested that exonization/pseudoexonization and indel were two main mechanisms to gain/loss and shift the miR156/529 binding sites after gene duplications, and the latter might easily introduce the mutation when miRNA binding sites shifted from the coding region to untranslated region. Therefore, this study provides the better understanding the functional diversity of SPL genes in land plants.

### List of abbreviations

| | |
|---|---|
| SBP | *Squamosa* promoter binding protein |
| UTR | untranslated region |
| CDS | coding sequences |
| TFs | transcription factors |
| *TSH4* | tasselsheath4 |
| *AP2* | *APETALA2* |
| SMART | Simple Modular Architecture Research Tool |
| indels | insertions and deletions |
| WGD | whole-genome duplication |
| PGDD | Plant Genome Duplication Database |
| Ks | synonymous substitution rate |
| Ka | nonsynonymous substitution rate |
| GSDS | Gene Structure Display Server |

## 1. Introduction

SQUAMOSA promoter binding protein-like (SPL) genes encode transcription factors (TFs) in plants. These genes share a highly conserved DNA-binding domain, termed the SBP domain (Cardon et al., 1999). This domain is an assembly of approximately 76 amino acid residues that are involved in both DNA binding and nuclear localization and features two zinc-binding sites (Klein et al., 1996; Yamasaki et al., 2004). The SPL proteins were originally isolated from *Antirrhinum majus* as factors that bind to the promoter region of the floral meristem identity gene *SQUAMOSA* (*SQUA*) (Klein et al., 1996). Since then, these genes have been found in various green plants, including single-celled green algae, mosses, gymnosperms, and angiosperms. Previous studies have suggested that land plant SPL genes originate from a common ancestor but diverge from algae SPL genes (Ling and Zhang, 2012a). It has been reported that land plant SPL genes are involved in the transitions from juvenile to adult phases and from vegetative to reproductive phases. Such functions of the SPL genes are elucidated by

* Corresponding author at: Department of Life Science, Liupanshui Normal University, Liupanshui, Guizhou 553004, China.
*E-mail address:* linglizhen@mail.kib.ac.cn (L.-Z. Ling).

https://doi.org/10.1016/j.plgene.2018.04.007
Received 15 September 2017; Received in revised form 30 March 2018; Accepted 30 April 2018
Available online 03 May 2018
2352-4073/ © 2018 Elsevier B.V. All rights reserved.

their interplay with *miR156* family members. For example, the low-level expression of SPL genes in an *miR156*-overexpression mutant prolonged the juvenile phase in maize (Chuck et al., 2007) and *Arabidopsis* (Wu and Poethig, 2006). Another miRNA gene, *miR529*, shares 14–16 nt of homology with *miR156* and targets the same SPL gene in plants (Ling and Zhang, 2012b). The best known example is *tasselsheath4* (*TSH4*) gene, which is involved in the bract development and the establishment of meristem boundaries in maize (Chuck et al., 2010). Land plant SPL genes are involved in trichome development, apical dominance, inflorescence branching, fruit ripening, plastochron length and pollen sac development (Preston and Hileman, 2013). In addition, an SPL gene, *SPL9*, has been reported to negatively regulate anthocyanin biosynthesis in the model plant *Arabidopsis* (Gou et al., 2011). Clearly, the SPL gene functions have diversified in land plants.

Gene duplication provides the raw genetic materials for the evolution of functionary novelty. Paralogous genes produced by gene duplication initially have identical sequences and functions but tend to diverge in both coding and regulatory regions and thereby acquire divergent expression and protein functions (Conant and Wagner, 2003; Kondrashov et al., 2002). Usually, divergence in regulatory regions can result in shifts in expression pattern, whereas changes in coding regions may lead to the acquisition of new functions due to amino acid substitutions (Gonzalez et al., 1995; Liu et al., 2015). Previous comprehensive phylogenomic analyses of sequenced plant genomes have elucidated that two groups of ancient genome duplication events (WGDs). One happened around 319 (ζ event) million years ago in the common ancestor of extant seed plants. The other happened about 192 (ε event) million years ago in the common ancestor of extant angiosperms (Jiao et al., 2011). Besides, most flowering plant lineages reflect one or more rounds genome duplication events. For example, the complete genome sequence of *Arabidopsis thaliana* has experienced two recent WGDs (named α and β) within the crucifer (Brassicaceae) lineage and one more ancient triplication (γ) shared with by all core eudicots (Barker et al., 2009; Tang et al., 2008; Vision et al., 2000). The extensive analyses of the *Populus trichocarpa* genome show evidence of the core eudicot triplication as well as a more recent WGD (Tuskan et al., 2006). Two polyploidy events in monocots (ρ and σ) have been inferred to have predated the diversification of cereal grains and other grasses (Poaceae) (Tang et al., 2010).

The previous study has described that the gene duplications have diversified the functions of SPL genes by sub-functionalization and neofunctionalization (Preston and Hileman, 2013). In addition, the phylogenetic and gene structural analysis have classed SPL genes of land plants into two groups (group I and II)and the latter has been further two subgroups (subgroup II-1 and II-2) (Zhang et al., 2015a). Interestingly, we found that the SPL genes targeted by *miR156* and *miR529* were only distributed in subgroup II-2, suggesting these targets have originated from the common ancestor. Meanwhile, we found that SPL genes in subgroup II-2 have experienced the largest copy number expansion when compared with those in other groups or subgroups (Zhang et al., 2015a). Although we have already compared the different evolutionary patterns of the SPL genes targeted by *miR156* and *miR529*, respectively (Zhang et al., 2015a and 2015b), yet much still remains to be learned about the evolutionary outcome of *miR156* and *miR529* targets after gene duplication. For example, how gene duplicate affects the binding sites of *miR156* and *miR529* on these SPL genes and whether it also affects their expression patterns and functions.

In the present study, we analyzed 136 SPL genes from eight land plants representing bryophytes, lycophytes, coniferopsida, basal angiosperms, monocots and eudicots. A total of 32 duplicated SPL target pairs were identified in these land plants, and several distinct evolutionary histories of duplication events were inferred in different lineages. Based on whether the miR156/529 binding sites were divergent, we first assorted all duplicated SPL target pairs into two classes: D-type (those with divergent miRNA binding sites) and S-type (those with the same miRNA binding sites). Meanwhile, there were included two divergent patterns: gain/loss of miRNA binding site (D-1 type) and mutation of miRNA binding site (D-2 type). We then analyzed and compared evolutionary rates, sequence identity and expression patterns of these paralogous genes among different types. In addition, we compared the structural changes between duplicated genes to analyze the divergent mechanism of miR156/529 binding sites in this study. Therefore, this study will help us to better understand the effect of gene duplication on functional diversity of SPL genes by diverging miR156/529 binding sites in land plants.

## 2. Materials and methods

### 2.1. Collection of SPL genes in land plants

The SPL gene candidates were obtained from gene prediction sets that were provided by Phytozome v10.3 (https://phytozome.jgi.doe.gov/pz/portal.html) (Goodstein et al., 2012). These prediction sets included seven representative land plants: one moss (*Physcomitrella patens*), one lycophyte (*Selaginella moellendorffii*), one base angiosperm (*Amborella trichopoda*), two eudicots (*Arabidopsis thaliana* and *Populus trichocarpa*) and two monocots (*Oryza sativa* subsp. *Japonica* and *Zea mays*). The corresponding protein, coding and genomic sequences were also downloaded from Phytozome v10.3 (https://phytozome.jgi.doe.gov/pz/portal.html) (Goodstein et al., 2012). For *Picea abies*, the SPL gene sequences were obtained from ConGenIE (http://congenie.org/) (Sundell et al., 2015). All SPL protein sequences were analyzed using SMART (Simple Modular Architecture Research Tool) (http://smart.embl-heidelberg.de/smart/set_mode.cgi?GENOMIC=1#) (Letunic et al., 2012) to confirm the presence of SBP domain. The genes containing partial SBP domains were excluded from further analyses due to the possibility of pseudogenes. Finally, a total of 136 SPL genes were determined from eight representative species of land plants. The SPL genes used in this study were summarized in Table S1.

### 2.2. Identification of duplicated SPL genes in land plants

Our previous study suggested that segmental and tandem gene duplications predominated during the expansion of the SPL gene family (Ling and Zhang, 2012a). Thus, we focused on the process of segmental and tandem duplication. Segmental duplications are usually defined as segments of DNA with near-identical sequences that map to two or more genomic locations. Segmental duplication occurs frequently in plants following polyploidy or whole-genome duplication (WGD), and numerous duplicated chromosomal blocks are strongly retained within genomes (Cannon et al., 2004). The Plant Genome Duplication Database (PGDD) (http://chibba.agtec.uga.edu/duplication/) (Lee et al., 2013) provides information on the synteny between chromosomes for 47 plants, including bryophytes and lycopodiophytes as well as angiosperms. Consequently, the synteny blocks of each SPL gene in one species were searched in PGDD to determine whether the SPL gene was involved in segmental duplication. Tandem duplications were characterized as multiple members of one family occurring within the same intergenic region or in neighboring intergenic regions (Maher et al., 2006). In this study, we defined adjacent homologous genes with no more than one intervening gene on the same chromosome as tandem duplicated genes.

### 2.3. Date estimation of the duplication events

The synonymous substitution (Ks) value of each gene pair within a segmental duplicated block was extracted from the PGDD dataset (http://chibba.agtec.uga.edu/duplication/) (Lee et al., 2013) and was used to calculate its mean Ks value. In dating segmental duplication events, the approximate time of the duplication event was evaluated using the mean Ks values.

Download English Version:

https://daneshyari.com/en/article/8647730

Download Persian Version:

https://daneshyari.com/article/8647730

Daneshyari.com