

RESEARCH

Open Access

# The *GRAS* gene family and its roles in seed development in litchi (*Litchi chinensis* Sonn)



Jingwen Chen<sup>1,2,3†</sup>, Qian Yan<sup>4†</sup>, Jiawei Li<sup>1,2,3</sup>, Lei Feng<sup>1,2,3</sup>, Yi Zhang<sup>1,2,3</sup>, Jing Xu<sup>1,2,3</sup>, Rui Xia<sup>1,2,3\*</sup>, Zaohai Zeng<sup>1,2,3\*</sup> and Yuanlong Liu<sup>1,2,3\*</sup>

## Abstract

**Background:** The *GRAS* gene family plays crucial roles in multiple biological processes of plant growth, including seed development, which is related to seedless traits of litchi (*Litchi chinensis* Sonn.). However, it hasn't been fully identified and analyzed in litchi, an economic fruit tree cultivated in subtropical regions.

**Results:** In this study, 48 LcGRAS proteins were identified and termed according to their chromosomal location. LcGRAS proteins can be categorized into 14 subfamilies through phylogenetic analysis. Gene structure and conserved domain analysis revealed that different subfamilies harbored various motif patterns, suggesting their functional diversity. Synteny analysis revealed that the expansion of the *GRAS* family in litchi may be driven by their tandem and segmental duplication. After comprehensively analysing degradome data, we found that four *LcGRAS* genes belong to HAM subfamily were regulated via miR171-mediated degradation. The various expression patterns of *LcGRAS* genes in different tissues uncovered they were involved in different biological processes. Moreover, the different temporal expression profiles of *LcGRAS* genes between abortive and bold seed indicated some of them were involved in maintaining the normal development of the seed.

**Conclusion:** Our study provides comprehensive analyses on *GRAS* family members in litchi, insight into a better understanding of the roles of *GRAS* in litchi development, and lays the foundation for further investigations on litchi seed development.

**Keywords:** Litchi, *GRAS* gene family, miR171, Seed development

## Background

*GRAS* is a major plant-specific transcription factor gene family among numerous transcription factors that are proved to function in plant growth and development, whose name is termed from the first three functionally characterized members, gibberellic-acid insensitive (GAI) [1], repressor of GAI (RGA) [2], and scarecrow (SCR) [3]. Typically, the *GRAS* proteins usually encompass 400–770

amino acids (aa) [4] and contain a variable N-terminal region and a highly conserved C-terminal region. The conserved region was composed of five motifs: LHRI, LHRIL, VHIID, PFYRE, and SAW [4–6]. *GRAS* proteins were divided into eight basic subfamilies in *Arabidopsis thaliana* and rice (*Oryza sativa* L.) according to their common feature [6]. Whereas in other plants such as *Prunus mume* [7], *Medicago truncatula* [8], Chinese Cabbage (*Brassica rapa* ssp. *pekinensis*) [9], pepper (*Capsicum annuum* L.) [10], tomato (*Solanum lycopersicum*) [11], sacred lotus (*Nelumbo nucifera*) [12], tea plant (*Camellia sinensis*) [13], *Populus* [14], pine (*Pinus radiata*) [15], castor beans (*Ricinus communis*) [16], Tartary buckwheat (*Fagopyrum tataricum*) [17], and cotton (*Gossypium hirsutum* L.) [18], the number of subfamilies varied from eight to 14,

\* Correspondence: [rxia@scau.edu.cn](mailto:rxia@scau.edu.cn); [zengzh@scau.edu.cn](mailto:zengzh@scau.edu.cn); [liyuanlong@scau.edu.cn](mailto:liyuanlong@scau.edu.cn)

†Jingwen Chen and Qian Yan contributed equally to this work.

<sup>1</sup>State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, South China Agricultural University, 483 Wushan Road, Tianhe, Guangzhou 510642, Guangdong Province, China  
Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

suggesting that specific subfamilies might be present in these species.

GRAS proteins are involved in many physiological processes such as signal transduction, root radial patterning and development, stress responses and meristem development [19]. However, each subfamily may have different functions. For instance, DELLA proteins act as inhibitory factors in the gibberellic acid (GA) signal transduction pathway and modulate the jasmonic acid (JA) signal transduction [20]. PAT1 is participating in phytochrome signaling of *Arabidopsis* [21]. Besides, SCR and SHR proteins are involved in root and shoot radial patterning in *Arabidopsis* [22]. SCL3 acts as an integrator downstream of the GA/DELLA and SCR/SHR pathways, mediating the GA-promoted cell elongation during root development [23]. Moreover, OsMOC1, AtLAS, and SILS play important roles in axillary meristem initiation, plant tillering, and control of grain yield [24–26]. Additionally, SCL13 with stress-related functions have been discovered in cabbage [27], petunia HAM mediates signals from differentiating cells for functioning in shoot meristem maintenance [28].

It has been mentioned in previous studies that some GRAS genes from HAM subfamily were targeted by miR171, which play diverse roles in plant development, such as flowering and phase transition [29, 30]. A miR171-SCL6 model contributes to embryogenic callus induction and torpedo-shaped embryo formation during somatic embryogenesis in two lily species [31]. In *Arabidopsis*, SCL6/SCL6-IV, SCL22/SCL6-III, and SCL27/SCL6-II are also known as targets of miR171, and play pivotal roles in the proliferation of meristematic cells [32, 33]. Overexpression of a tomato miR171 target gene, SIGRAS24, impacts multiple agronomical traits such as plant height, flowering time, root length, fruit set and development [34].

Litchi is a subtropical fruit tree of family Sapindaceae with great economic and nutritional value. Based on fruit anatomy, the fruit of litchi is a drupe with an edible aril enclosing a single seed surrounded by a pericarp [35]. Small seeds or seedlessness is an economically desirable trait of litchi, which lead to great market value. In flowering plants, seed development is preceded by a double fertilization event, which form the precursor cells of embryo and endosperm. The endosperm is essential for the development of an embryo and the rapid disintegration of endosperm lead to the abortion of seed. Coordination in the growth of endosperm and embryo is crucial during early seed development, results in the discrepancy in seed size in a fruit [36]. A bunch of GRAS genes were supposed to take part in determining endosperm and embryo development in several species, affecting seed development. In lily, a GRAS gene belonging to the LISCL subfamily, plays a role in the

microsporogenesis process of the anther [37]. GS6, a unique member of the GRAS gene family, was responsible for the reduction of grain size and weight during the domestication of rice [38]. In *P. mume*, 25% of GRAS genes showed higher expression in seeds [7]. In apple, the higher expression level of *MdGRAS126*, *MdGRAS18*, and *MdGRAS79* in seeds in accordance with the finding in *P. mume* [7, 39], indicating the important roles of GRAS genes in seed development. However, the biological function of GRAS proteins in seed development of litchi remains scarce. Hence, a comprehensive analysis of GRAS genes in litchi would be informative in laying foundation for the characterization of their potential function, especially in seed development.

In our study, two different litchi cultivars ('Huaizhi' and 'NMC') and a wildtype litchi ('WL10') were used to carry out a genome-wide analysis of GRAS genes in litchi. As a result, 48 *LcGRAS* genes were identified. Gene structure, phylogeny, chromosomal distributions, duplication events, dual synteny analysis, and miRNA-mediated regulation were characterized. Expression pattern of *LcGRAS* genes was detected in various tissues and four different stages during seed development, including 15, 25, 35, and 45 days after anthesis (DAA), representing near globular-shaped, heart-shaped, torpedo-shaped and cotyledon-shaped embryo stage. Among them, *LcGRAS1*, *LcGRAS15*, *LcGRAS24*, *LcGRAS28*, *LcGRAS29*, *LcGRAS40*, and *LcGRAS48* were found to exert its potential function in seed development of litchi via auxin and GA pathway.

## Results

### Identification and phylogenetic analysis of *LcGRAS* genes

Based on homology analysis, 48 *LcGRAS* proteins were identified from the litchi genome (Additional file 1: Table S1, Additional file 2), which were renamed from *LcGRAS1* to *LcGRAS48* according to the chromosomal location. The length of *LcGRAS* proteins was between 422 aa (*LcGRAS47*) and 803 aa (*LcGRAS34*). The predicted molecular weight (MW) of the proteins ranged from 47.24 kDa (*LcGRAS18*) to 89.89 kDa (*LcGRAS34*), and the predicted isoelectric point (pI) ranged from 4.62 (*LcGRAS30*) to 8.67 (*LcGRAS7*) (Additional file 1: Table S2). The numerical range of the above characteristics is similar to that of other species [7, 10], indicating that our identification of *LcGRASs* is relatively accurate, and the basic characteristics of GRASs are relatively conserved in different species.

To explore the phylogenetic relationship of *LcGRAS* protein, we constructed a phylogenetic tree based on the amino acid sequences of 48 *LcGRAS*, 32 *AtGRAS*, and 53 *OsGRAS* proteins. According to previous studies [6, 18], 48 *LcGRAS* members were divided into 14 subfamilies: SCR, SHR, DELLA, PAT1, HAM, LISCL, LAS,

SCL3, SCL4/7, DLT, Os4, Os19, Os43 and L\_GRAS (Fig. 1). The HAM subfamily possessed the most LcGRAS members (11), followed by LISCL (9), PAT1 (6), SHR (5), DELLA (3), SCR (3), SCL3 (1), LAS (1), AtSCL4/7 (1), and DLT (1). In addition, 2, 1, 1 of LcGRAS proteins were respectively grouped into the Os4, Os19, and Os43 subfamilies, all of which were previously reported as rice-specific [14]. These LcGRAS members may exist before the divergence of dicotyledons and monocotyledons and lost in *Arabidopsis*. Furthermore, the L\_GRAS subfamily contained three members, all of which were from litchi, implying that this litchi-specific subfamily may have unique functions in litchi or close species.

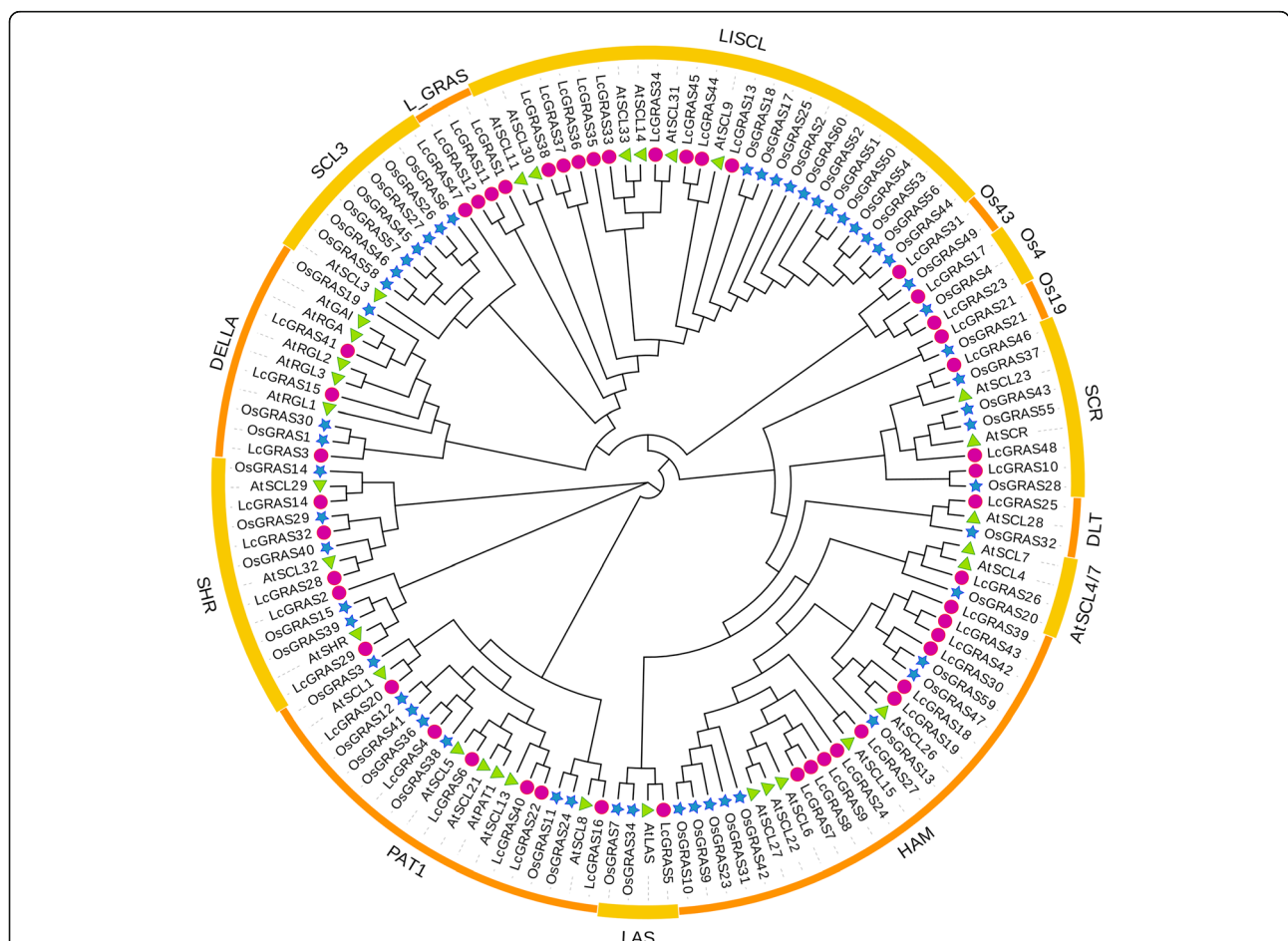
**Gene structure and conserved domain analysis**

To further understand the composition of *LcGRASs*, the gene structures of them were compared. 81.3% of the *LcGRASs* were intronless, only nine *LcGRAS* members had one or more introns (Fig. 2a, b). All *LcGRAS*

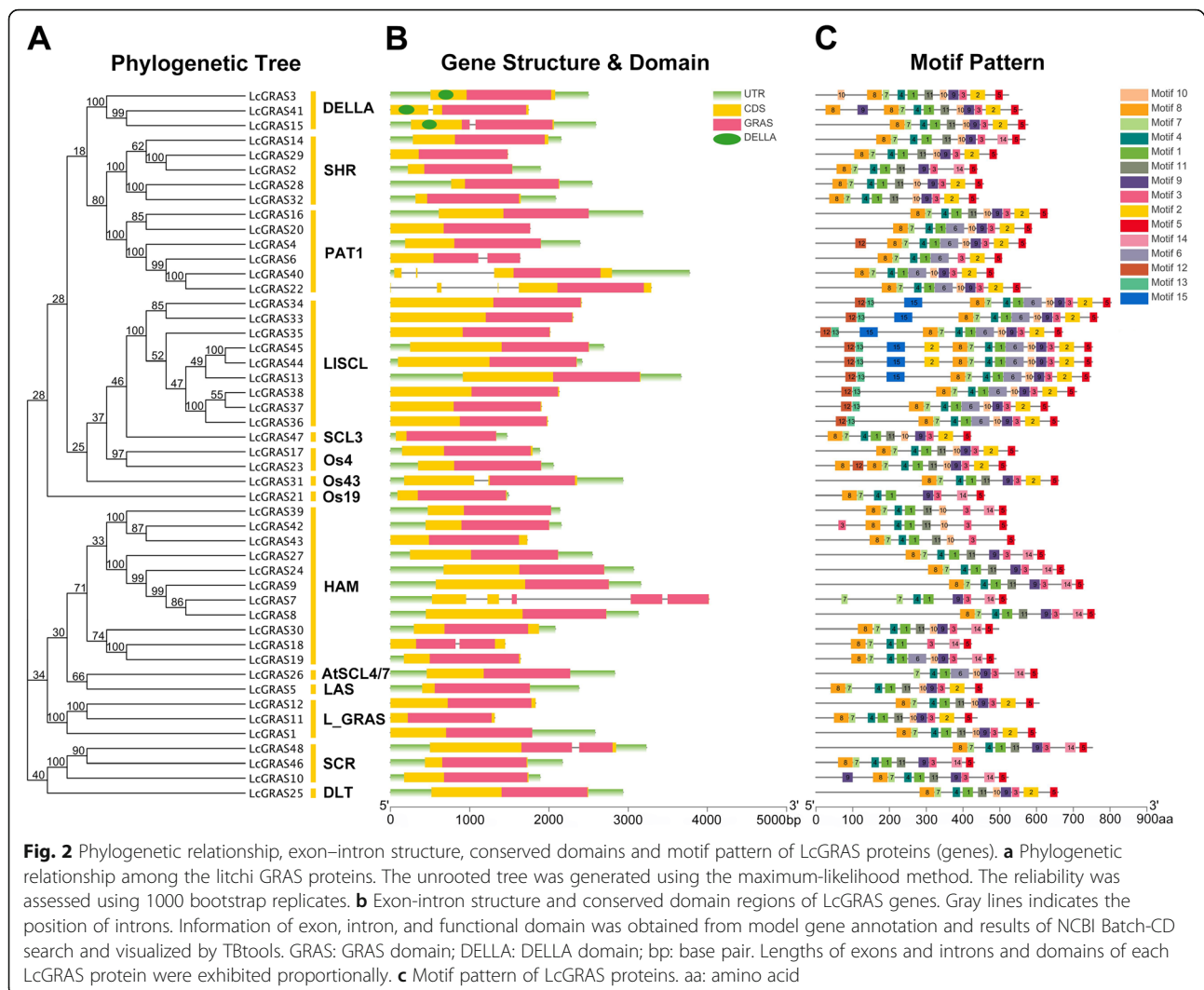
proteins incorporated the GRAS domain. Among them, three members (LcGRAS3, LcGRAS15, LcGRAS41) occupied a DELLA domain, which was essential for GA signal perception. A total of 15 distinct conserved motifs (named motif 1–15) were identified in our motif analysis and almost all LcGRAS proteins contain motif 1, 3, 4, 5, 7, 8, 9, 10, 11, 14 (Fig. 2a, c). Interestingly, motif locations exhibited subfamily specific patterns. For example, motif 6 only existed in PAT1, LISCL, HAM, and AtSCL4/7 subfamilies, while motif 12, 13, and 15 were only located in the N-terminal of the members in LISCL subfamily (Fig. 2a, c). In general, different subfamilies embraced various structure compositions, suggesting their great functional diversity.

**Chromosomal distribution and synteny analysis of *LcGRAS* genes**

*LcGRAS* genes were unevenly distributed on the 15 chromosomes of litchi (Fig. 3). There are 7 *LcGRAS* genes located in Chr3, Chr13, and Chr15 respectively,



**Fig. 1** Phylogenetic tree of the *GRAS* gene family in litchi, rice, and *Arabidopsis*. The phylogenetic tree was constructed using Maximum-Likelihood (ML) method by MEGA7.0. Subfamilies were marked by bold yellow lines or orange lines in the external circle. The pink circles, green triangles and blue stars represent *GRAS* genes from litchi, *Arabidopsis* and rice, respectively

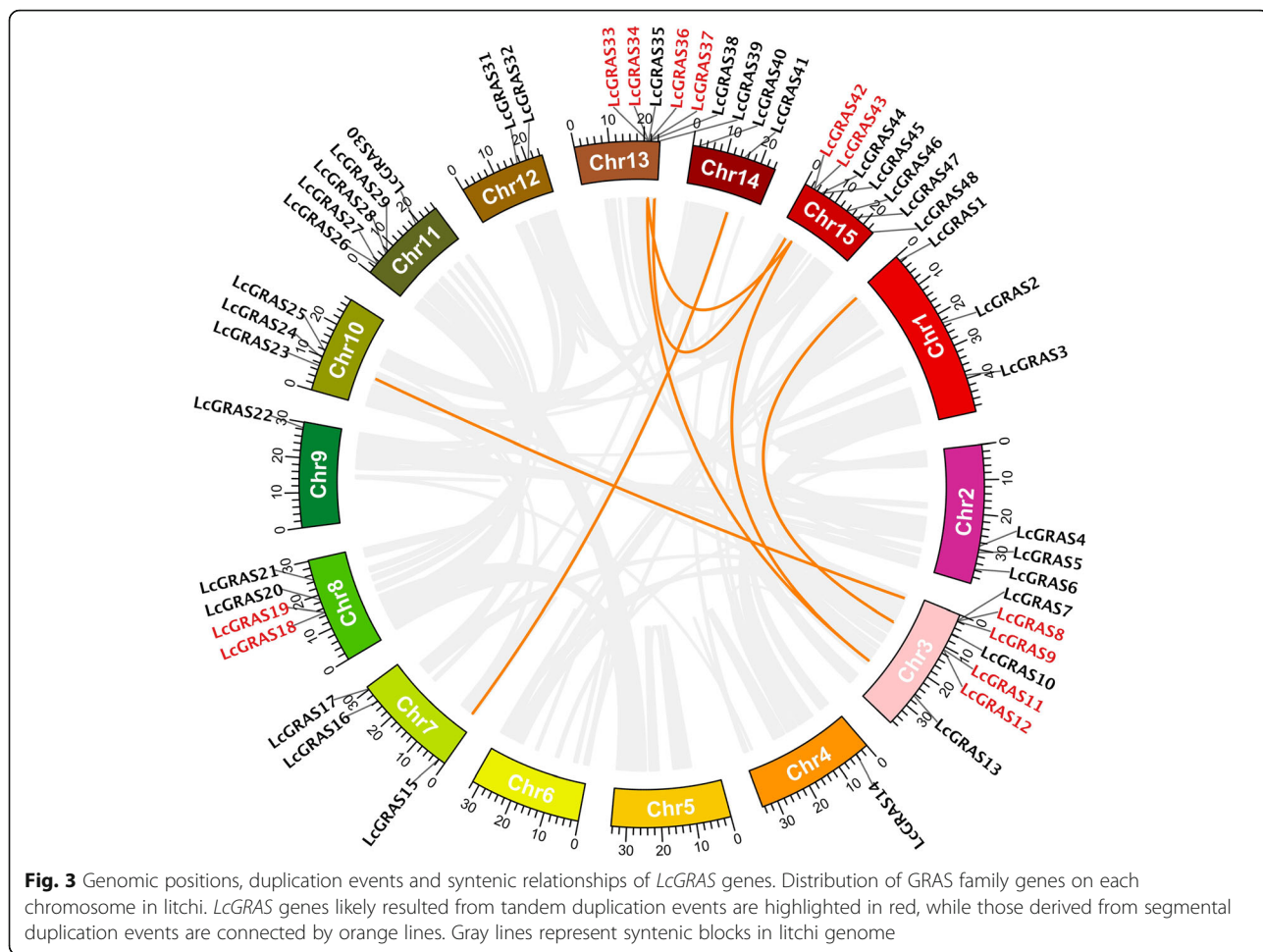


followed by 5 in Chr11 and 4 in Chr8. Four chromosomes (Chr1, Chr2, Chr7, Chr10) had three *LcGRAS* loci, while two of them (Chr12, Chr14) contained two *LcGRAS* genes, respectively. Chr4 and Chr9 possessed only one *LcGRAS* loci (Fig. 3). Gene duplication contributed to the amplification of the *LcGRAS* family. Tandem duplication (highlighted in red in Fig. 3) was presented in Chr3, Chr8, Chr13 and Chr15, indicating that they were hot spots for *LcGRAS* gene distributions (Fig. 3). Moreover, seven pairs of segmental duplication genes (orange lines) were detected between chromosomes: Chr1/Chr3, Chr3/Chr10, Chr3/Chr15, Chr3/Chr13, Chr7/Chr14 and Chr13/Chr15 (2 pairs) (Fig. 3). Tandem and segmental duplication events of *LcGRAS* genes occurred mainly in HAM, LISCL, DELLA, and L\_GRAS subfamilies. Taken together, these results suggested that tandem and segmental duplication may have been the main driving force of the evolution of the litchi *GRAS* family.

To further deduce the relationship of the *LcGRAS* genes, we checked their synteny with *GRAS* genes from a dicyledonous plant (*A. thaliana*) and one monocyledonous plant (*O. sativa*) (Fig. 4). A total of 13 *LcGRAS* genes showed syntenic relationships with *AtGRAS* genes, and 8 of them had syntenic loci in rice. *LcGRAS28*, *LcGRAS29*, *LcGRAS33*, and *LcGRAS44* had syntenic loci in both *Arabidopsis* and rice (Fig. 4, Additional file 1: Table S3). These four genes belonged to either the SHR (*LcGRAS28*, *LcGRAS29*) or the LISCL (*LcGRAS33*, *LcGRAS44*) subfamily, hinting their conserved biological function in plants.

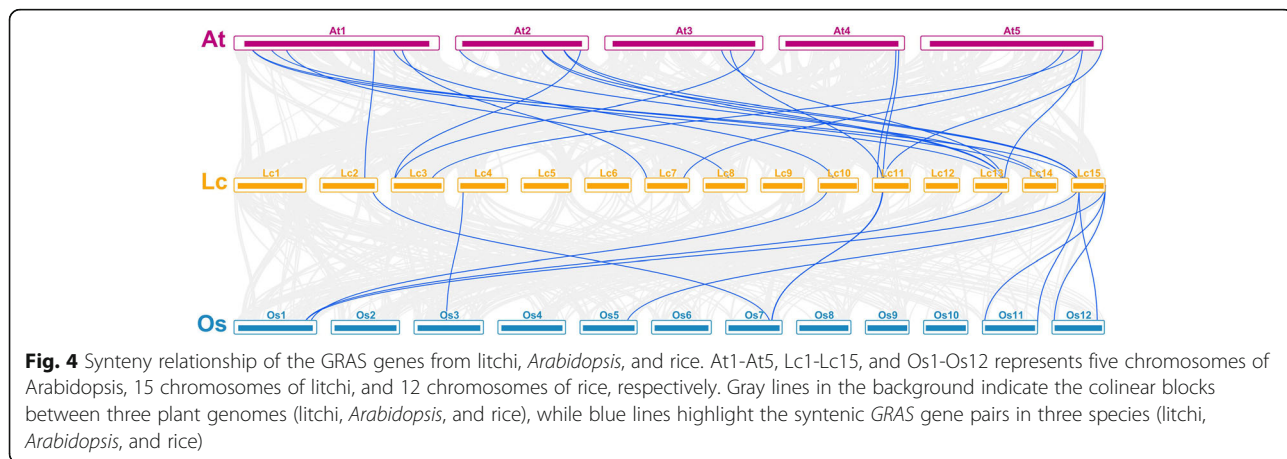
#### Analyses of miRNA targeting *LcGRAS* genes

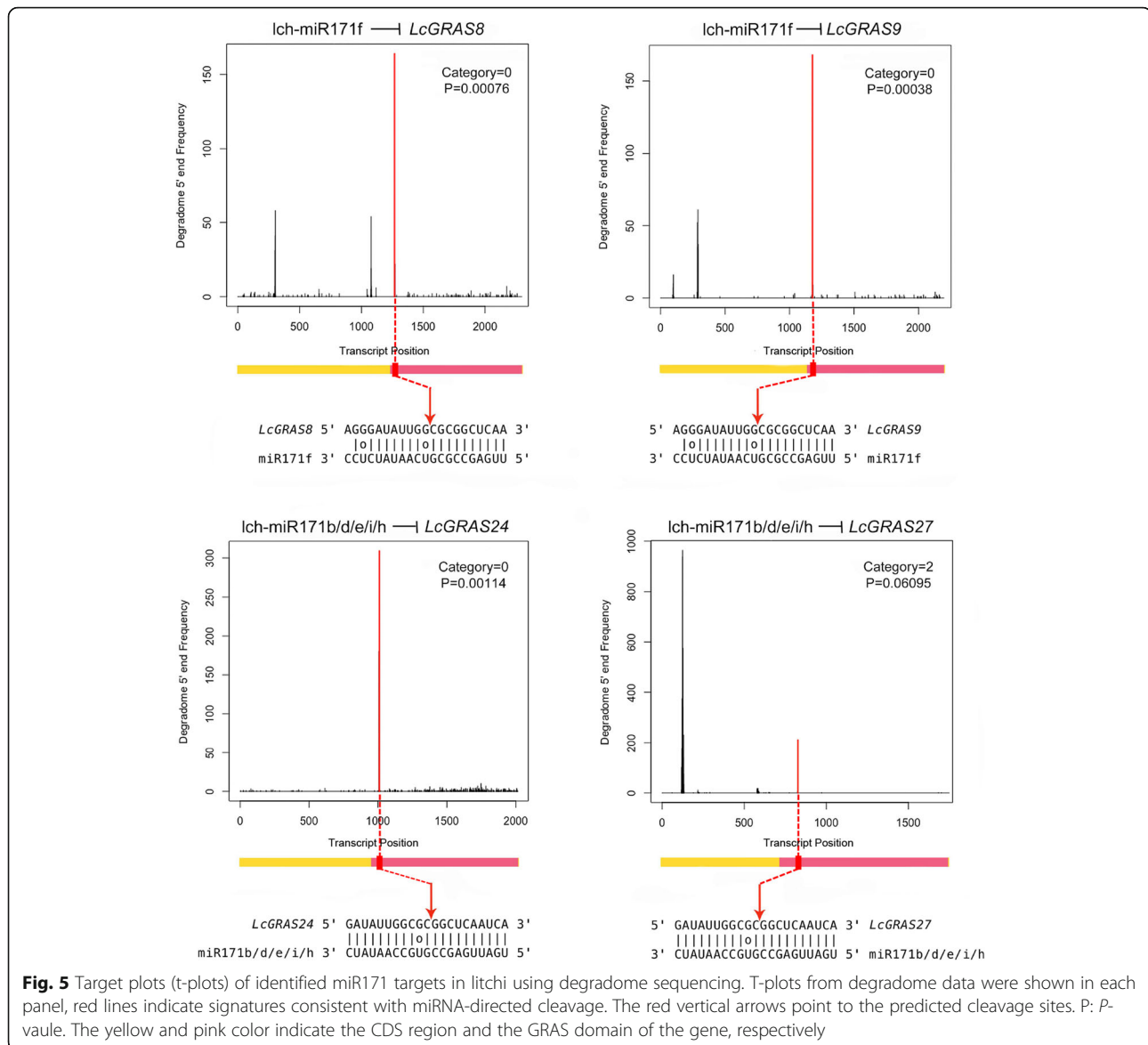
microRNAs are crucial regulatory factors in plants. They regulate the expression of target genes at post-transcriptional level [40]. By combining analyses of degradome data sets from four different libraries, eight miRNAs might have the potential to regulate *LcGRAS*



genes (Additional file 1: Table S4). Target genes with penalty score less than 5 and category less than 2 were considered confident [41]. Thus, in total four members (*LcGRAS8*, *LcGRAS9*, *LcGRAS24*, *LcGRAS27*) of *LcGRAS* genes were identified as targets of miR171 (Fig. 5, Additional file 1: Table S4). The miR171-mediated cleavages

were verified using degradome data, as presented in the form of target plots (t-plots), showing the abundance of cleaved tags relative to their positions in the transcripts. For each miR171 targeted *LcGRAS* genes, a clear cleavage was detected at the target site of lch-miR171s (Fig. 5). All miR171 target sites were located at the fore end



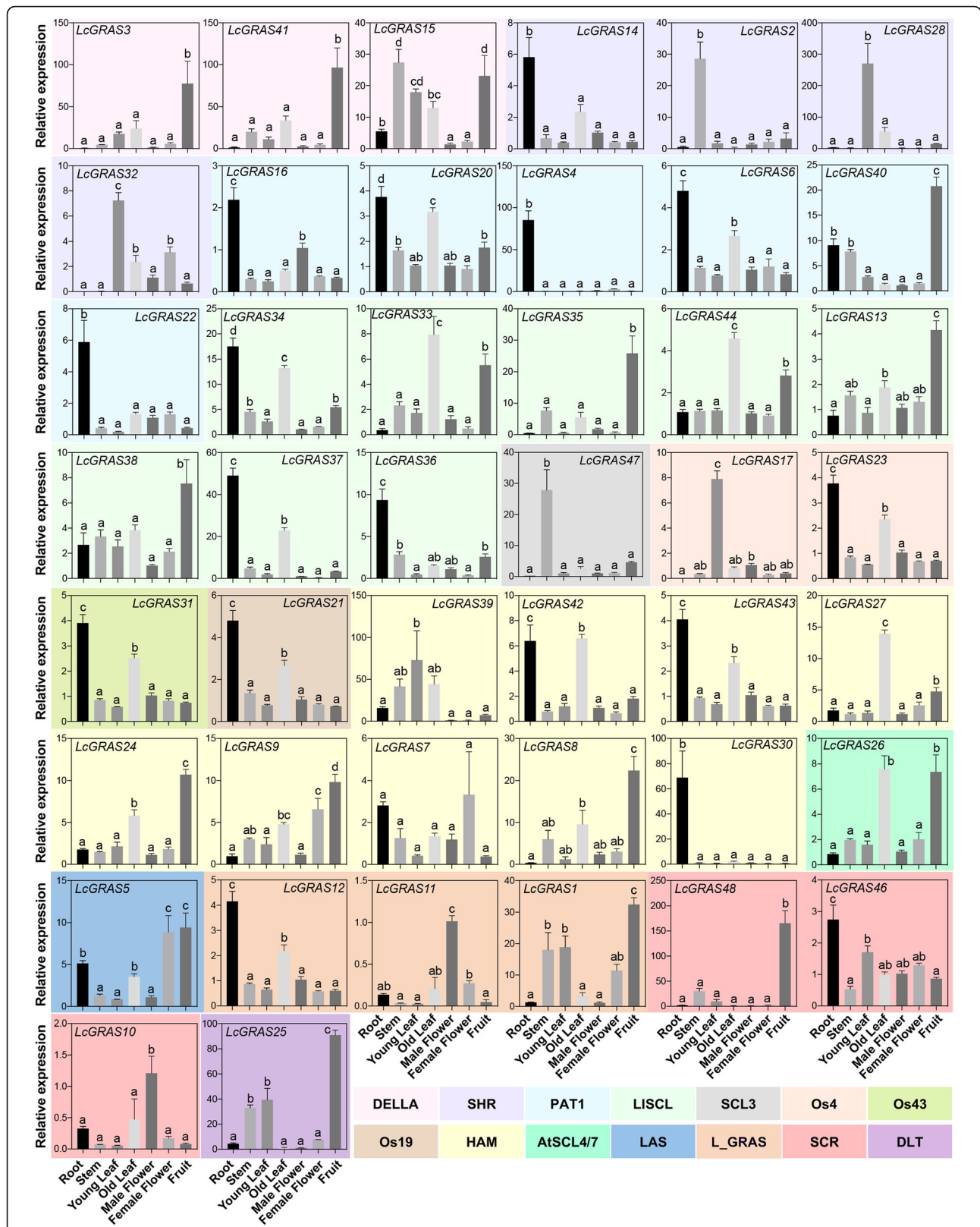


of the GRAS domain (Fig. 5), and all miR171 targeted *LcGRAS*s belonged to the HAM subfamily, which were supposed to function in meristematic cell development, root length, and flowering [42–44].

#### Expression analysis of *LcGRAS* genes in different tissues

To investigate the role of these *LcGRAS* genes, RT-qPCR was used to analyze the expression pattern of 48 *LcGRAS* genes in seven tissues, including root, stem, young leaf, old leaf, male flower, female flower, and fruit. As illustrated in Fig. 6, expression of 44 *LcGRAS* genes was obtained, while four *LcGRAS* genes cannot be detected because of their extremely low expression levels. Most of the *LcGRAS* genes were highly expressed in root and old leaf and poorly expressed in either male or female flowers. In addition, most genes in *LISCL* were

highly expressed in fruit and old leaf, while some genes were with rich expression in root, suggesting that functional diversification was present in this subfamily. *LcGRAS46* in SCR, *LcGRAS14* in SHR and almost all genes in PAT1 subfamily were abundant in root (Fig. 6); genes in DELLA subfamily (*LcGRAS3*, *LcGRAS15*, *LcGRAS41*) were all highly expressed in fruit; *LcGRAS2* (SHR) and *LcGRAS47* (SCL3) were found to have higher expression in stem (Fig. 6). We also found that gene *LcGRAS11* which belongs to litchi-specific subfamily L-GRAS, was highly expressed in male flower (Fig. 6), demonstrating it would be closely related to the male flower development. In contrast, *LcGRAS5*, the sole member of the LAS subfamily, was highly expressed in female flower and fruit (Fig. 6), indicating its function in the development of female flowers and fruits. Moreover,



**Fig. 6** Relative expression of LcGRASs in seven tissues. Data from three independent biological replicates are shown with standard error (SE). Different letters above the bars represent significant differences ( $P < 0.05$ , LSD) among seven tissues. The control used in the calculations was male flower. The same background color represents members of the same subfamily

*LcGRAS25* (DLT) was of high-level expression in fruit, stem and young leaf (Fig. 6), implying its potential function in fruit, stem, and young leaf development.

#### Expression profile of *LcGRAS* genes in two varieties with contrasting seed size

To explore the character of *LcGRAS*s in the regulation of seed development in litchi, transcriptome analysis was conducted on a litchi cultivar ('NMC') and a wildtype litchi ('WL10'), representing developing small (abortive) and large (bold) seeds, respectively (Fig. 7a, Additional file 1: Table S5). Seed samples at four developmental stages, including globular-shaped embryo stage (15 DAA), heart-shaped embryo stage (25 DAA), torpedo-shaped embryo stage (35 DAA), and cotyledon-shaped embryo stage (45 DAA), were collected for RNA sequencing. Pairwise comparison of the developing seeds unmasked the common and exclusive differentially expressed transcripts at 15, 25, 35, and 45 DAA between the two varieties. Among 48 *LcGRAS* genes, 8, 12, 10 and 7 *LcGRAS* genes were differentially expressed ('NMC' vs 'WL10') at 15, 25, 35, and 45 DAA, respectively (Fig. 7b, c, Additional file 1: Table S6). Notably, *LcGRAS32* was consistently high expressed in abortive-seeded cultivar ('NMC') during all four stages (Fig. 7b, c), which suggested its potential function in embryo abortive development, giving rise to small seed. The seed development process in litchi could be divided into the cell division stage and the filling stage around 28 DAA when the embryo reached the heart-shaped embryo stage with a rudimentary cotyledon. The cell division stage before 28 DAA was more important for normal seed development [45]. Intriguingly, *LcGRAS29* and *LcGRAS40* were exclusively high accumulated in small-seed during both globular-shaped embryo stage (15 DAA) and heart-shaped embryo stage (25 DAA), implying their possible function in seed abortion. In addition, *LcGRAS1* was specifically highly expressed in globular-shaped embryo stage (15 DAA) while *LcGRAS24* was specifically highly expressed in the heart-shaped embryo stage (25 DAA) (Fig. 7c), indicating that they may be involved in cell division in seed development as well. Furthermore, *LcGRAS15*, *LcGRAS28* and *LcGRAS48* were of high expression in 'WL10' (bold-seeded) during torpedo- and cotyledon-shaped embryo stages, suggesting these genes may be linked with important traits during the filling stage of normal seed development (Fig. 7c).

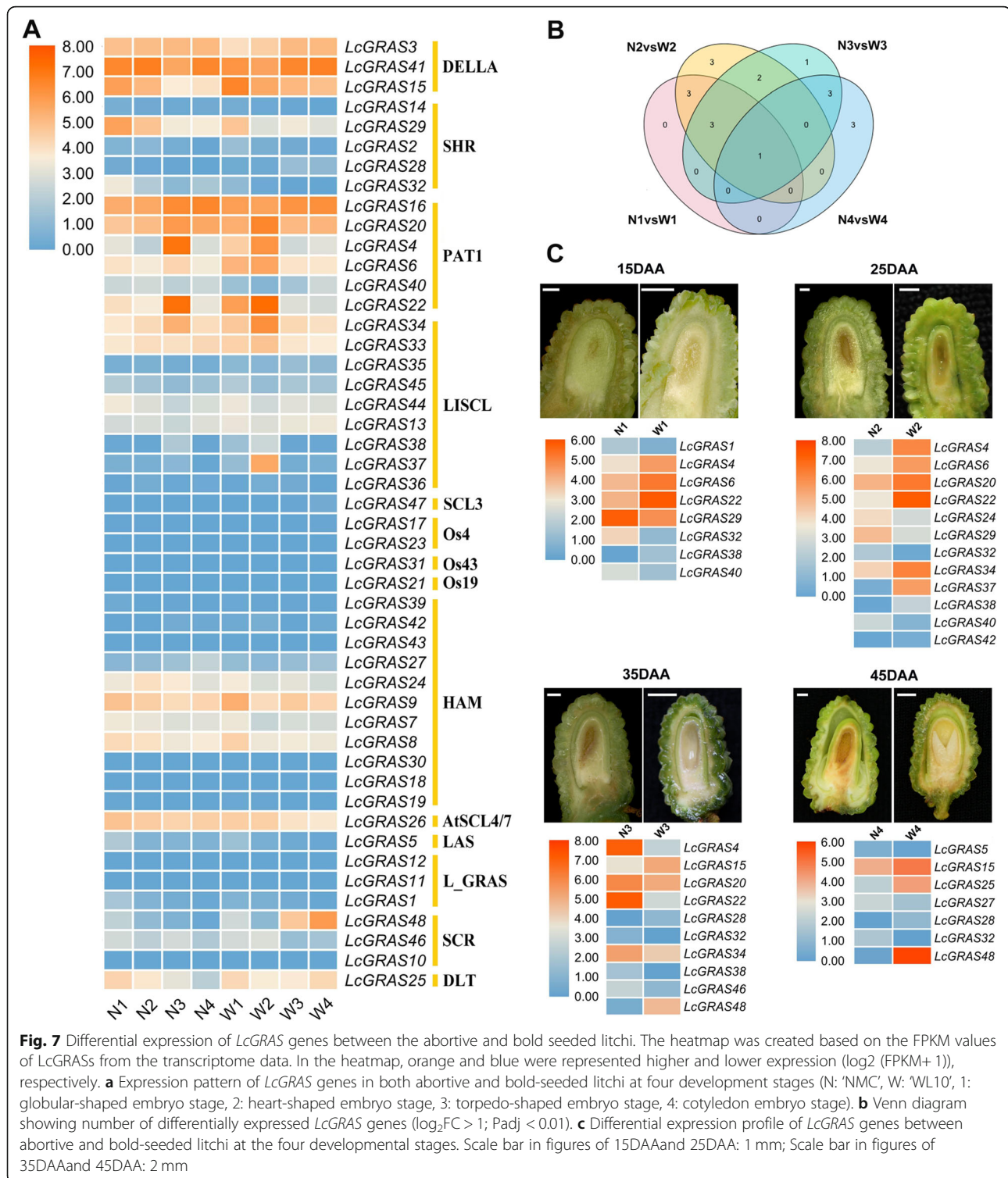
#### Discussion

*GRAS* gene family has been characterized in several plant species, and involved in numerous critical development and physiological processes. In our study, 48 *LcGRAS* genes were identified in litchi (Fig. 1). The population of *LcGRAS* members was larger than that in

*Arabidopsis* (33) [46], and roughly the same to that in *P. mume* (46) [7], tomato (*S. lycopersicum*) (53) [11], castor bean (*R. communis*) (48) [16], and pepper (*C. annuum* L.) (50) [10], but less than in *Populus* (106) [14], rice (57) [6], and cotton (*Gossypium hirsutum* L.) (150) [18], implying extensive duplication and diversification of the *LcGRAS* gene family among species. In the analysis of the structural compositions of *LcGRAS*s, we found 81.3% of *LcGRAS*s were intronless (Fig. 2), which was similar to tomato (77.4%) [11] and *P. mume* (82.2%) [7]. Intronless genes have been discovered in gene families DEAD box RNA helicase [47] and F-box gene family [48]. The high proportion of intronless genes in litchi suggests that they may have experienced intron loss events during evolution, which is common in other eukaryotes [49]. Tandem and segmental duplication are the main mechanisms for the expansion of plant gene families [50] and play a crucial role in the adaptive response to environmental stimuli [51]. Tian and colleagues [6] analyzed the expansion mechanism of *GRAS* gene families in *Arabidopsis* and rice. In litchi, six tandem duplication and seven segmental duplication gene pairs were found (Fig. 3), these gene pairs in HAM, LISCL, L\_GRAS and DELLA subfamily are like other plant species [52]. Our result indicated that duplication events may be a mechanism for expanding the number of *GRAS*s in these subfamilies. In addition, *LcGRAS1*, *LcGRAS11*, and *LcGRAS12* of L\_GRAS subfamily, a new subfamily identified in litchi, were also experienced the two types of duplication events. These L\_GRAS genes may possess special function for the growth and development of litchi, which need further researches. Moreover, some gene pairs of the two duplication events had similar expression pattern in different tissues, as shown in Fig. 6, which also existed between *PmGRAS16* and *PmGRAS26* in *P. mume* [7].

The analysis of expression patterns can facilitate our depiction of the potential functions of *GRAS* genes [53, 54]. Genes in maize (*ZmSCR*) and rice (*OsSCR*) were shown to have similar expression patterns to *AtSCR* in roots [55, 56], and these two *GRAS* members (*SCR* and *SHR*), were involved in several different stages of (root) development [57]. *GRAS* protein from the PAT1 clade was shown to be associated with the development of the adventitious and lateral root [58, 59]. HAM clade of the *GRAS* family was vital for root development but involved in leaf development, with a triple-mutant (*scl6*, *scl22*, and *scl27*) leading to reduced root growth and abnormal leaf patterning [44, 60]. In our study, four *LcGRAS* genes were undetectable in all tissues, suggesting a trend to degenerate these genes after gene duplication or the loss of gene functions during evolution. Most of the *LcGRAS* genes were highly expressed in root, including members in *SCR*, *SHR*, *PAT1*, and *HAM*





subfamily, indicating they might function in root development in litchi. Moreover, GRAS protein from the HAM subfamily was supposed to participate in the vegetative to the reproductive phase transition by activating the miR156-SPLs pathway [42, 61]. *LcGRAS8*, *LcGRAS9*,

*LcGRAS24*, *LcGRAS27* from HAM subfamily were identified as the targets of miR171 (Fig. 5), and they shared a similar expression pattern, suggested that the relatively conserved functions of the miR171-GRASs regulatory networks in litchi. The four genes especially highly

expressed in both old leaf and fruit, which implied their potential roles in leaf patterning, flower organ formation [42] and fruit development in litchi [62–64]. Additionally, several genes were prominently with higher expression in fruit, including three DELLA genes (*LcGRAS3*, *LcGRAS15*, *LcGRAS41*), and *LcGRAS25* in DLT subfamily, which might be involved in the fruit development through GA signal transduction pathway [65–68] or brassinosteroid signal transduction pathway [69, 70], respectively.

GA and auxin are prominently associated with the seed formation during fruit development [71–73]. Moreover, some GRAS proteins function as regulators of auxin and GA in plant development, such as fruit and seed development. For instance, the overexpression of *SIGRAS7* enhancing GA/auxin signaling and improving resistance to abiotic stresses [74]. *SIGRAS24* was characterized to impact multiple agronomical traits by regulating auxin and GA homeostasis in tomato [34]. Additionally, it is evident that *SIGRAS40* acted as a regulator of auxin and GA as the overexpression of *SIGRAS40* led to auxin insensitivity and GA deficiency [75]. Moreover, overexpressing *SIGRAS24* or *SIGRAS40* in plants would lead to pleiotropic phenotypes such as reduced fruit set ratio, arrested fruit, and abnormal seed development [34, 75]. *SIGRAS24* and *SIGRAS40* were the target genes of miR171 and belonged to HAM subfamily in tomato. Similarly, in litchi, there were 4 genes of HAM subfamily targeted by miR171, which were higher expressed in fruit, indicating miR171-GRAS regulatory pathway might play similar roles in seed and fruit development like *SIGRAS24* and *SIGRAS40* through GA signal transduction pathway. Expect that, in the *SIDE*LLA deficit model, the tomatoes exhibited GA insensitivity and displayed a GA-constitutive response phenotype, including parthenocarpy [66, 68]. In our study, three DELLA genes (*LcGRAS3*, *LcGRAS15*, and *LcGRAS41*) were higher expressed in fruit, indicating their potential functions in fruit development through GA pathway. *PrSCL1* (*Pinus radiata* SCL1) and *CsSCL1* (*Castanea sativa* SCL1) were shown to regulate adventitious root formation through auxin signaling [76]. In *Arabidopsis*, the collaboration between the SHR-SCR complex and auxin influx carriers (LAX3 and AUX1) could lead to synergistic effect on primary/lateral root development [77]. In pine and cucumber, relatively high expressing GRAS transcripts, such as SCR and SHR, were measured in non-differentiated proliferating embryogenic cultures and during embryo development [15, 78]. In our study, *LcGRAS14* in SHR was highly expressed in root, and *LcGRAS48* was highly expressed in fruit, indicating that *LcGRAS* members might function in root and seed development by participating in auxin signal pathway. Hence, based on their similar expression

pattern in different tissues and similar conserved domain, *LcGRASs* that belong to HAM, DELLA and SHR were supposed to be involved in different developmental processes via crosstalk with GA or auxin signaling.

‘NMC’ and ‘WL10’ are two litchi varieties that display remarkable difference in seed size after maturity. ‘WL10’ produces larger seeds with normally developed embryos and cotyledons, while ‘NMC’, as the seed-aborting cultivar, produces seeds with defect embryos or cotyledons. It has been reported that 28 DAA represented a transition point between the cell division stage and the filling stage during litchi seed development, after which sequential liquid endosperm and embryo development were not observed in ‘NMC’ [45]. In our result, *LcGRAS29* (SHR), *LcGRAS40* (PAT1), *LcGRAS1* (LGRAS) and *LcGRAS24* (HAM) were exclusively and highly accumulated in ‘NMC’ (abortive seed) before 28 DDA (Fig. 7c), indicating its exceptional function in endosperm and embryo abortion of litchi. In addition, *LcGRAS15* (DELLA), *LcGRAS28* (SHR) and *LcGRAS48* (SCR) were up-regulated in ‘WL10’ (bold-seed) at 35DAA and 45DAA (Fig. 7c), which suggested that they may work in later seed maturation. These genes might participate in seed development by regulating auxin and GA pathways.

## Conclusions

In this study, 48 *LcGRAS* genes were identified in litchi and divided into 14 subfamilies. Members of the same subfamily have similar gene structures. Some *LcGRAS* genes are derived from gene duplication. The expression patterns of *LcGRAS* genes in different tissues were diverse, indicating that they might have different functions during the development of litchi. Four *LcGRAS* genes were regulated by miR171 directly. In addition, our result indicated *LcGRAS* genes are differentially expressed in different varieties of litchi (‘NMC’ and ‘WL10’) and illustrated crucial roles of *LcGRAS* proteins in embryos or cotyledons development which affects seed size. This research was the first comprehensive identification of *LcGRAS* genes in litchi. These results provide the foundation to elucidate the regulation mechanism of *LcGRASs* in plant growth and seed size, showing that *LcGRASs* might have important functions in litchi breeding.

## Methods

### Plant materials preparation

Three 13-year-old ‘Huaizhi’ litchi (one of the main cultivars in China) trees used in our study were planted in the orchard located at South China Agricultural University (Guangzhou, China). Different tissues, including root (root tips approximately 10 cm long), stem, young leaves (leaves approximately 3 cm long with yellow or light

green color, and the tip of the leaves is red), mature leaves (green but not leathery leaves), male flower (full bloom), female flower (full bloom), and young fruit (31 DAA) were collected for RT-qPCR analyses. The 30-year-old 'NMC' (cultivar litchi with abortive seeds) and 'WL10' (wildtype litchi with bold seeds) used in this study were grown in the germplasm resource orchard of Guangdong Province Fruit Research Institution (Guangzhou, China). Seed samples of 'NMC' and 'WL10' used in RNA-seq analysis were collected in different developmental stages (15, 25, 35, and 45 DAA) from two randomly selected trees. All samples were collected separately from three trees with similar growing conditions, and then quickly frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$ .

#### Identification and protein property analysis of LcGRASs

The Gtf/ Gff3 Sequence extractor in TBtools V1.046 [79] was used to extract the coding sequences (CDS) of all GRAS genes from a reference litchi ('Feizixiao') genome of 15 pseudo-chromosomes (470 Mb) with 96.2% completeness (assembled in house, data unpublished yet) based on the gene structure annotation information, and then CDS sequences were translated into protein sequences using Batch Translate CDS to protein tool in TBtools V1.046 [79]. Thirty-three GRAS protein sequences from *Arabidopsis* were downloaded from TAIR ([https://www.arabidopsis.org/browse/genefamily/gras\\_genefamily.jsp](https://www.arabidopsis.org/browse/genefamily/gras_genefamily.jsp)) [5], which were used as baits to identify potential GRAS genes in the litchi genome by BLAST analysis with a relative sensitive cutoff (E-value set at  $1e-5$ ) in TBtools V1.046 [79], the resultant protein sequences were then used as queries to search against the UniportKB/Swiss-port (swissport) databases using the BLASTP program with default parameters to avoid false positives. The identified sequences were then validated using CDD (<http://www.ncbi.nlm.nih.gov/cdd/>) [80] with E-value threshold 0.01 and Pfam (<http://pfam.xfam.org/>) [81] databases with default parameters. The ProtParam tools from the ExPASy website (<https://web.expasy.org/prot-param/>) [82] were used to obtain the sequence length, predicted molecular weight, and predicted isoelectric point of the identified GRAS proteins.

#### Phylogenetic analysis of LcGRASs, AtGRASs, and OsGRASs

GRAS genes of *Arabidopsis thaliana* and rice together with the litchi GRAS genes were used in phylogenetic analysis. Thirty-three GRAS proteins from *Arabidopsis thaliana* and 60 GRAS proteins from rice (Additional file 1: Table S1) were downloaded from TAIR [5] and PlantTFDB V5.0 (<http://planttfdb.cbi.pku.edu.cn/>) [83] respectively. AtSCL16 (a putative pseudogene [46]) and some members in rice (protein length is less than 350 aa [11]) were excluded in the subsequent analysis. Multiple

protein sequence alignment was carried out via Muscle [84], and the poorly aligned regions were removed by TrimAL 1.3 in TBtools V1.046 with default parameters [79]. Phylogenetic analysis was performed by MEGA7.0 program by maximum likelihood (ML) method and the bootstrap test was carried out with 1000 iterations [85].

#### Gene structure, domain, and conserved motif analysis

Introns and exons of each LcGRAS gene were analyzed using TBtools V1.046 [79]. The conserved domains were defined using the Batch-CD search (<http://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>) [86, 87] with default parameters. The MEME-Suite 5.1.1 online program (<http://meme-suite.org/>) [88] was used to analyze the conserved motifs to investigate the structural differences among LcGRAS members. All results above were visualized by TBtools V1.046 [79].

#### Chromosomal distribution and gene duplication of LcGRAS genes

The physical location information was obtained from the litchi gff3 file and plotted by TBtools V1.046 [79]. Multiple collinear scanning toolkits (MCScanX) with default parameters were used to analyze gene duplication events [89]. The syntenic relationship between *LcGRASs*, *AtGRASs*, and *OsGRASs* was determined using MCScanX and visualized by multiple synteny plot tool in TBtools V1.046 [79].

#### Identification of transcripts targeted by miRNAs

Six sRNA and four degradome data sets of litchi (*L. chinensis* Sonn.) were downloaded from accession number GSE98698 which were stored in NCBI [90]. An in-house software, sRNAMiner, was used to monitor quality, trim adaptor, and collapse reads with the same sequence of sRNA sequencing data [91]. Subsequently, noncoding RNAs (including rRNA, snoRNA, and tRNA) and sRNAs from chloroplast and mitochondrial genome were removed by mapping against RNA Family (Rfam) database V13.0 [92, 93] and the Plant organelles database [94] via bowtie [95] respectively. Preprocessed reads were mapped to the litchi genome and used to explore miRNAs. Cleveland 4.0 was adopted to identify transcripts targeted by miRNAs and authentic target sites with a confident level of category 0–2 and penalty score no more than 5 [96] were screened. All degradome reads on cleave sites were normalized to reads per 10 million (RPTM).

#### Expression analysis of LcGRAS genes by RT-qPCR

Total RNA was extracted using the Hot borate method described by Wan and Wilkins [97], the cDNA strand was synthesized with the HiScriptII Q RT SuperMix for qPCR (+gDNA wiper) (Vazyme Cat No. R223-01). RT-qPCR was performed with GoTaq<sup>®</sup> qPCR and RT-qPCR

Systems (Promega Cat No. A6001) using a Light Cycler 480 Real-Time PCR Detection System (Roche, Rotkreuz, Switzerland). Primers of *LcGRASs*, and two reference genes *GAPDH* and *EF* [98] were designed by Primer Premier 5.0 (Additional file 1: Table S7). Each expression profile was independently verified in three biological replicates. The relative expression level of each gene was calculated by the  $2^{-\Delta\Delta Ct}$  method [99].

### RNA-seq and differential expression analysis

The transcriptomic data were generated from different seed development stages (15, 25, 35, and 45 DAA) of two species of litchi ('NMC' and 'WL10'). Trimmomatic software was used to control the quality of raw RNA-seq data and remove the adapter [100]. Afterwards STAR software was used to map clean data to the litchi genome and the expression level of transcripts was normalized into fragments per kilobase of transcript per million fragments mapped (FPKM) by StringTie [101, 102]. Differentially expressed genes were identified using an R package, DESeq2 [103], where adjusted *P*-value (Padj) < 0.01 and foldchange > 2 were set as thresholds. In detail, we took the average of the two biological replicate counts of each sample, and then divided the average counts of all the two sets of samples to be compared with each other to get the fold change value (FC). Lastly, took the logarithm of 2 for the obtained fold change, next got log2FC (log2 fold change). If the log2FC value of a gene was greater than 1, and the Padj was less than 0.01, the gene would be significantly up-regulated. Correspondingly, if the log2FC value was lower than - 1, and the Padj was less than 0.01, the gene was considered to be significantly down-regulated.

### Abbreviations

GAI: Gibberellic-acid insensitive; RGA: Repressor of GAI; SCR: Scarecrow; aa: Amino acids; GA: Gibberellic acid; JA: Jasmonic acid; DAA: Day after anthesis; MW: Molecular weight; pl: Isoelectric point; T-plots: Target plots; CDS: Coding sequences; RPTM: Reads per ten million; FPKM: Fragments per kilobase of transcript per million fragments mapped; Padj: Adjusted *P*-value; FC: Fold change

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12870-021-03193-1>.

**Additional file 1: Table S1.** GRAS proteins of litchi, *Arabidopsis*, and rice. **Table S2.** Information of *LcGRAS* genes. **Table S3.** Syntenic gene pairs among litchi, *Arabidopsis* and rice. **Table S4.** Information of miRNA targets in litchi GRAS gene family. **Table S5.** Gene expression profile of 48 *LcGRAS* members among four seed development stages (Normalized as FPKM). **Table S6.** Information of differentially expressed *LcGRAS* genes. **Table S7.** Specific primers of 48 *LcGRAS* genes used for qPCR in this study.

**Additional file 2.** Sequences of litchi GRAS genes.

### Acknowledgements

Not applicable.

### Authors' contributions

J. C., Q. Y., R. X., Z. Z. and Y. L. conceived and designed the experiments. J. C. performed the experiments, prepared the materials and the manuscript. Q. Y. conducted the RNA-seq of 'NMC' and 'WL10'. J. C., J. L. and L. F. participated in data analyses. Y. L., J. C., Z. Z., Y. Z., J. X. and R. X. revised the manuscript. All authors have read and approved the final manuscript.

### Funding

This work was supported by funds from the National Natural Science Foundation of China (grant NO. 31872063 and NO. 32002009) for the design and analysis of the data, the Basic and Applied Basic Research Programs of Foundation of Guangdong Province (2019A1515110595) for the RNA sequencing, and the Outstanding Talent Program of the Ministry of Agriculture and the China Agricultural Research System (201804020063) provide experimental platform.

### Availability of data and materials

Six sRNA and four degradome data sets of litchi (*L. chinensis* Sonn.) were available from accession number GSE98698 which were stored in NCBI. The litchi genome data and RNA-seq data that support the findings of this study have been deposited into CNGB Sequence Archive (CNSA, <https://db.cngb.org/cnsa/>) of China National GeneBank DataBase (CNGBdb) with accession number CNP0001024 and CNP0001865 respectively, which will be released after the publication of the related paper. Review links are available as <https://db.cngb.org/cnsa/project/CNP0001024/reviewlink/> and <http://db.cngb.org/cnsa/project/CNP0001865/reviewlink/> respectively. Other data sets supporting the results of this article are included within the article and its additional files.

### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, South China Agricultural University, 483 Wushan Road, Tianhe, Guangzhou 510642, Guangdong Province, China. <sup>2</sup>Key Laboratory of Biology and Germplasm Enhancement of Horticultural Crops in South China, Ministry of Agriculture, South China Agricultural University, Guangzhou, China. <sup>3</sup>Guangdong Litchi Engineering Research Center, College of Horticulture, South China Agricultural University, Guangzhou, China. <sup>4</sup>Key Laboratory of South Subtropical Fruit Biology and Genetic Resource Utilization, Ministry of Agriculture / Guangdong Province Key Laboratory of Tropical and Subtropical Fruit Tree Research / Institute of Fruit Tree Research, Guangdong Academy of Agricultural Sciences, Guangzhou, China.

Received: 18 May 2021 Accepted: 25 July 2021

Published online: 17 September 2021

### References

- Peng J, Carol P, Richards DE, King KE, Cowling RJ, Murphy GP, et al. The *Arabidopsis* GAI gene defines a signaling pathway that negatively regulates gibberellin responses. *Genes Dev.* 1997;11(23):3194–205. <https://doi.org/10.1101/gad.11.23.3194>.
- Silverstone AL, Ciampaglio CN, Sun T. The *Arabidopsis* RGA gene encodes a transcriptional regulator repressing the gibberellin signal transduction pathway. *Plant Cell.* 1998;10(2):155–69. <https://doi.org/10.1105/tpc.10.2.155>.
- Laurenzio LD, Wysocka-Diller J, Malamy JE, Pysh L, Helariutta Y, Freshour G, et al. The SCARECROW gene regulates an asymmetric cell division that is essential for generating the radial organization of the *Arabidopsis* root. *Cell.* 1996;86(3):423–33. [https://doi.org/10.1016/S0092-8674\(00\)80115-4](https://doi.org/10.1016/S0092-8674(00)80115-4).
- Pysh LD, Wysocka-Diller JW, Camilleri C, Bouchez D, Benfey PN. The GRAS gene family in *Arabidopsis*: sequence characterization and basic expression

- analysis of the *SCARECROW-LIKE* genes. *Plant J.* 1999;18(1):111–9. <https://doi.org/10.1046/j.1365-3113.1999.00431.x>.
5. Bolle C. The role of GRAS proteins in plant signal transduction and development. *Planta.* 2004;218(5):683–92. <https://doi.org/10.1007/s00425-004-1203-z>.
  6. Tian C, Wan P, Sun S, Li J, Chen M. Genome-wide analysis of the GRAS gene family in rice and *Arabidopsis*. *Plant Mol Biol.* 2004;54(4):519–32. <https://doi.org/10.1023/B:PLAN.0000038256.89809.57>.
  7. Lu J, Wang T, Xu Z, Sun L, Zhang Q. Genome-wide analysis of the GRAS gene family in *Prunus mume*. *Mol Gen Genomics.* 2014;290(1):303–17. <https://doi.org/10.1007/s00438-014-0918-1>.
  8. Zhang H, Cao Y, Shang C, Li J, Wang J, Wu Z, et al. Genome-wide characterization of GRAS family genes in *Medicago truncatula* reveals their evolutionary dynamics and functional diversification. *PLoS One.* 2017;12(9):e0185439. <https://doi.org/10.1371/journal.pone.0185439>.
  9. Song X, Liu T, Duan W, Ma Q, Ren J, Wang Z, et al. Genome-wide analysis of the GRAS gene family in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *Genomics.* 2014;103(1):135–46. <https://doi.org/10.1016/j.ygeno.2013.12.004>.
  10. Liu B, Sun Y, Xue J, Jia X, Li R. Genome-wide characterization and expression analysis of GRAS gene family in pepper (*Capsicum annuum* L.). *PeerJ.* 2018;6:e4796.
  11. Huang W, Xian Z, Kang X, Tang N, Li Z. Genome-wide identification, phylogeny and expression analysis of GRAS gene family in tomato. *BMC Plant Biol.* 2015;15(1):209. <https://doi.org/10.1186/s12870-015-0590-6>.
  12. Wang Y, Shi S, Zhou Y, Zhou Y, Yang J, Tang X. Genome-wide identification and characterization of GRAS transcription factors in sacred lotus (*Nelumbo nucifera*). *PeerJ.* 2016;4:e2388. <https://doi.org/10.7717/peerj.2388>.
  13. Wang Y, Liu Z, Wu Z, Li H, Wang W, Cui X, et al. Genome-wide identification and expression analysis of GRAS family transcription factors in tea plant (*Camellia sinensis*). *Sci Rep.* 2018;8(1):3949. <https://doi.org/10.1038/s41598-018-22275-z>.
  14. Liu X, Widmer A. Genome-wide comparative analysis of the GRAS gene family in *Populus*, *Arabidopsis* and Rice. *Plant Mol Biol Rep.* 2014;32(6):1129–45. <https://doi.org/10.1007/s11105-014-0721-5>.
  15. Abarca D, Pizarro A, Hernández I, Sánchez C, Solana SP, Amo AD, et al. The GRAS gene family in pine: transcript expression patterns associated with the maturation-related decline of competence to form adventitious roots. *BMC Plant Biol.* 2014;14(1):354. <https://doi.org/10.1186/s12870-014-0354-8>.
  16. Xu W, Chen Z, Ahmed H, Han B, Cui Q, Liu A. Genome-wide identification, evolutionary analysis, and stress responses of the GRAS gene family in castor beans. *Int J Mol Sci.* 2016;17(7):1004. <https://doi.org/10.3390/ijms17071004>.
  17. Liu M, Huang L, Ma Z, Sun W, Wu Q, Tang Z, et al. Genome-wide identification, expression analysis and functional study of the GRAS gene family in Tartary buckwheat (*Fagopyrum tataricum*). *BMC Plant Biol.* 2019;19(1):342. <https://doi.org/10.1186/s12870-019-1951-3>.
  18. Zhang B, Liu J, Yang Z, Chen E, Zhang C, Zhang X, et al. Genome-wide analysis of GRAS transcription factor gene family in *Gossypium hirsutum* L. *BMC Genomics.* 2018;19(1):348. <https://doi.org/10.1186/s12864-018-4722-x>.
  19. Mayrose M, Ekengren SK, Melech-Bonfil S, Martin GB, Sessa G. A novel link between tomato GRAS genes, plant disease resistance and mechanical stress response. *Mol Plant Pathol.* 2006;7(6):593–604. <https://doi.org/10.1111/j.1364-3703.2006.00364.x>.
  20. Hou X, Lee LYC, Xia K, Yan Y, Yu H. DELLAs modulate jasmonate signaling via competitive binding to JAZs. *Dev Cell.* 2010;19(6):884–94. <https://doi.org/10.1016/j.devcel.2010.10.024>.
  21. Bolle C, Koncz C, Chua NH. PAT1, a new member of the GRAS family, is involved in phytochrome A signal transduction. *Genes Dev.* 2000;14(10):1269–78.
  22. Helariutta Y, Fukaki H, Wysocka-Diller J, Nakajima K, Jung J, Sena G, et al. The *SHORT-ROOT* gene controls radial patterning of the *Arabidopsis* root through radial signaling. *Cell.* 2000;101(5):555–67. [https://doi.org/10.1016/S0092-8674\(00\)80865-X](https://doi.org/10.1016/S0092-8674(00)80865-X).
  23. Zhang Z, Ogawa M, Fleet CM, Zentella R, Hu J, Heo JO, et al. *SCARECROW-LIKE 3* promotes gibberellin signaling by antagonizing master growth repressor *DELLA* in *Arabidopsis*. *Proc Natl Acad Sci U S A.* 2011;108(5):2160–5. <https://doi.org/10.1073/pnas.1012232108>.
  24. Li X, Qian Q, Fu Z, Wang Y, Xiong G, Zeng D, et al. Control of tillering in rice. *Nature.* 2003;422(6932):618–21. <https://doi.org/10.1038/nature01518>.
  25. Greb T, Clarenz O, Schäfer E, Müller D, Herrero R, Schmitz G, et al. Molecular analysis of the *LATERAL SUPPRESSOR* gene in *Arabidopsis* reveals a conserved control mechanism for axillary meristem formation. *Genes Dev.* 2003;17(9):1175–87. <https://doi.org/10.1101/gad.260703>.
  26. Schumacher K, Schmitt T, Rossberg M, Schmitz G, Theres K. The *Lateral suppressor (Ls)* gene of tomato encodes a new member of the VHIID protein family. *Proc Natl Acad Sci U S A.* 1999;96(1):290–5. <https://doi.org/10.1073/pnas.96.1.290>.
  27. Park HJ, Jung WY, Lee SS, Song JH, Kwon S-Y, Kim H, et al. Use of heat stress responsive gene expression levels for early selection of heat tolerant cabbage (*Brassica oleracea* L.). *Int J Mol Sci.* 2013;14(6):11871–94. <https://doi.org/10.3390/ijms140611871>.
  28. Stuurman J, Jaggi F, Kuhlemeier C. Shoot meristem maintenance is controlled by a GRAS-gene mediated signal from differentiating cells. *Genes Dev.* 2002;16(17):2213–8. <https://doi.org/10.1101/gad.230702>.
  29. Jover-Gil S, Candelà H, Ponce M-R. Plant microRNAs and development. *Int J Dev Biol.* 2005;49(5-6):733–44. <https://doi.org/10.1387/ijdb.05210515j>.
  30. Fattash I, Vo B, Reski R, Hess WR, Frank W. Evidence for the rapid expansion of microRNA-mediated regulation in early land plant evolution. *BMC Plant Biol.* 2007;7(1):13. <https://doi.org/10.1186/1471-2229-7-13>.
  31. Li H, Zhang J, Yang Y, Jia N, Wang C, Sun H. miR171 and its target gene *SCL6* contribute to embryogenic callus induction and torpedo-shaped embryo formation during somatic embryogenesis in two lily species. *Plant Cell Tiss Org.* 2017;130:591–600.
  32. Schulze S, Schafer BN, Parizotto EA, Voinnet O, Theres K. *LOST MERISTEMS* genes regulate cell differentiation of central zone descendants in *Arabidopsis* shoot meristems. *Plant J.* 2010;64(4):668–78. <https://doi.org/10.1111/j.1365-3113.2010.04359.x>.
  33. Engstrom EM, Andersen CM, Gumalak-Smith J, Hu J, Orlova E, Sozzani R, et al. *Arabidopsis* homologs of the petunia *HAIRY MERISTEM* gene are required for maintenance of shoot and root indeterminacy. *Plant Physiol.* 2011;155(2):735–50. <https://doi.org/10.1104/pp.110.168757>.
  34. Huang W, Peng S, Xian Z, Lin D, Hu G, Yang L, et al. Overexpression of a tomato miR171 target gene *SIGRAS24* impacts multiple agronomical traits via regulating gibberellin and auxin homeostasis. *Plant Biotechnol J.* 2016;15(4):472–88. <https://doi.org/10.1111/pbi.12646>.
  35. Wang H, Lai B, Huang X. Litchi fruit set, development, and maturation. In: *The lychee biotechnology*. Singapore: Springer; 2017. p. 1–30. [https://doi.org/10.1007/978-981-10-3644-6\\_1](https://doi.org/10.1007/978-981-10-3644-6_1).
  36. Chaudhury AM, Koltunow A, Payne T, Luo M, Tucker MR, Dennis ES, et al. Control of early seed development. *Annu Rev Cell Dev Biol.* 2001;17(1):677–99. <https://doi.org/10.1146/annurev.cellbio.17.1.677>.
  37. Morohashi K, Minami M, Takase H, Hotta Y, Hiratsuka K. Isolation and characterization of a novel GRAS gene that regulates meiosis-associated gene expression. *J Biol Chem.* 2003;278(23):20865–73. <https://doi.org/10.1074/jbc.M301712200>.
  38. Sun L, Li X, Fu Y, Zhu Z, Tan L, Liu F, et al. *GS6*, a member of the GRAS gene family, negatively regulates grain size in rice. *J Integr Plant Biol.* 2013;55(10):938–49. <https://doi.org/10.1111/jipb.12062>.
  39. Fan S, Zhang D, Gao C, Zhao M, Wu H, Li Y, et al. Identification, classification, and expression analysis of GRAS gene family in *Malus domestica*. *Front Physiol.* 2017;8:253. <https://doi.org/10.3389/fphys.2017.00253>.
  40. German MA, Pillay M, Jeong D-H, Hetawal A, Luo S, Janardhanan P, et al. Global identification of microRNA–target RNA pairs by parallel analysis of RNA ends. *Nat Biotechnol.* 2008;26(8):941–6. <https://doi.org/10.1038/nbt1417>.
  41. Feng L, Xia R, Liu Y. Comprehensive characterization of miRNA and PHAS loci in the diploid strawberry (*Fragaria vesca*) genome. *Hortic Plant J.* 2019;5(6):255–67. <https://doi.org/10.1016/j.hpj.2019.11.004>.
  42. Curaba J, Talbot M, Li Z, Helliwell C. Over-expression of microRNA171 affects phase transitions and floral meristem determinacy in barley. *BMC Plant Biol.* 2013;13(1):6. <https://doi.org/10.1186/1471-2229-13-6>.
  43. Fan T, Li X, Yang W, Xia K, Ouyang J, Zhang M. Rice Osa-miR171c mediates phase change from vegetative to reproductive development and shoot apical meristem maintenance by repressing four *OsHAM* transcription factors. *PLoS One.* 2015;10(5):e0125833. <https://doi.org/10.1371/journal.pone.0125833>.
  44. Wang L, Mai Y, Zhang Y, Luo Q, Yang H. MicroRNA171c-targeted *SCL6-II*, *SCL6-III*, and *SCL6-IV* genes regulate shoot branching in *Arabidopsis*. *Mol Plant Pathol.* 2010;3(5):794–806. <https://doi.org/10.1093/mp/ssp042>.
  45. Zhang J, Wu Z, Hu F, Liu L, Huang X, Zhao J, et al. Aberrant seed development in *Litchi chinensis* is associated with the impaired expression of cell wall invertase genes. *Hortic res.* 2018;5(1):39. <https://doi.org/10.1038/s41438-018-0042-1>.
  46. Lee M-H, Kim B, Song S-K, Heo J-O, Yu N-I, Lee SA, et al. Large-scale analysis of the GRAS gene family in *Arabidopsis thaliana*. *Plant Mol Biol.* 2008;67(6):659–70. <https://doi.org/10.1007/s11103-008-9345-1>.

47. Aubourg S, Kreis M, Lecharny A. The *DEAD box RNA* helicase family in *Arabidopsis thaliana*. *Nucleic Acids Res.* 1999;27(2):628–36. <https://doi.org/10.1093/nar/27.2.628>.
48. Jain M, Nijhawan A, Arora R, Agarwal P, Ray S, Sharma P, et al. F-box proteins in rice. Genome-wide analysis, classification, temporal and spatial gene expression during panicle and seed development, and regulation by light and abiotic stress. *Plant Physiol.* 2007;143(4):1467–83. <https://doi.org/10.1104/pp.106.091900>.
49. Rogozin IB, Carmel L, Csuros M, Koonin EV. Origin and evolution of spliceosomal introns. *Biol Direct.* 2012;7(1):11. <https://doi.org/10.1186/1745-6150-7-11>.
50. Cannon SB, Mitra A, Baumgarten A, Young ND, May G. The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol.* 2004;4(1):10. <https://doi.org/10.1186/1471-2229-4-10>.
51. Hanada K, Zou C, Lehti-Shiu MD, Shinozaki K, Shiu SH. Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. *Plant Physiol.* 2008;148(2):993–1003. <https://doi.org/10.1104/pp.108.122457>.
52. Wu N, Zhu Y, Song W, Li Y, Yan Y, Hu Y. Unusual tandem expansion and positive selection in subgroups of the plant *GRAS* transcription factor superfamily. *BMC Plant Biol.* 2014;14(1):373. <https://doi.org/10.1186/s12870-014-0373-5>.
53. Shan Z, Luo X, Wu M, Wei L, Fan Z, Zhu Y. Genome-wide identification and expression of *GRAS* gene family members in cassava. *BMC Plant Biol.* 2020;20(1):46. <https://doi.org/10.1186/s12870-020-2242-8>.
54. Wang L, Ding X, Gao Y, Yang S. Genome-wide identification and characterization of *GRAS* genes in soybean (*Glycine max*). *BMC Plant Biol.* 2020;20(1):415. <https://doi.org/10.1186/s12870-020-02636-5>.
55. Lim J, Helariutta Y, Specht CD, Jung J, Sims L, Bruce WB, et al. Molecular analysis of the *SCARECROW* gene in maize reveals a common basis for radial patterning in diverse meristems. *Plant Cell.* 2002;12(8):1307–18. <https://doi.org/10.1105/tpc.12.8.1307>.
56. Kamiya N, Itoh J-I, Morikami A, Nagato Y, Matsuoka M. The *SCARECROW* gene's role in asymmetric cell division in rice plants. *Plant J.* 2003;36(1):45–54. <https://doi.org/10.1046/j.1365-3113.2003.01856.x>.
57. Bolle C. Functional aspects of *GRAS* family proteins. In: Gonzalez DH, editor. *Plant transcription factors*. NewYork: Academic; 2016. p. 295–311.
58. Cai H, Xuan L, Xu L, Huang M, Xu M. Identification and characterization of nine *PAT1* branch genes in poplar. *Plant Growth Regul.* 2016;81:1–10.
59. Battaglia M, Ripodas C, Clua J, Baudin M, Aguilar OM, Niebel A, et al. A nuclear factor Y interacting protein of the *GRAS* family is required for nodule organogenesis, infection thread progression, and lateral root growth. *Plant Physiol.* 2014;164(3):1430–42. <https://doi.org/10.1104/pp.113.230896>.
60. Ma Z, Hu X, Cai W, Huang W, Zhou X, Luo Q, et al. *Arabidopsis* miR171-targeted scarecrow-like proteins bind to GT cis-elements and mediate gibberellin-regulated chlorophyll biosynthesis under light conditions. *PLoS Genet.* 2014;10(8):e1004519. <https://doi.org/10.1371/journal.pgen.1004519>.
61. Xue X, Zhao B, Chao L, Chen D, Cui W, Mao Y, et al. Interaction between two timing microRNAs controls trichome distribution in *Arabidopsis*. *PLoS Genet.* 2014;10(4):e1004266. <https://doi.org/10.1371/journal.pgen.1004266>.
62. Wu X, Liu M, Ge X, Xu Q, Guo W. Stage and tissue-specific modulation of ten conserved miRNAs and their targets during somatic embryogenesis of Valencia sweet orange. *Planta.* 2011;233(3):495–505. <https://doi.org/10.1007/s00425-010-1312-9>.
63. Zhai L, Xu L, Wang Y, Huang D, Yu R, Limerica C, et al. Genome-wide identification of embryogenesis-associated microRNAs in radish (*Raphanus sativus* L.) by high-throughput sequencing. *Plant Mol Biol Rep.* 2014;32:900–15.
64. Zhang S, Jian Z, Han S, Yang W, Li W, Wei H, et al. Four abiotic stress-induced miRNA families differentially regulated in the embryogenic and non-embryogenic callus tissues of *Larix leptolepis*. *Biochem Biophys Res Commun.* 2010;398(3):355–60. <https://doi.org/10.1016/j.bbrc.2010.06.056>.
65. Arnaud N, Girin T, Sorefan K, Fuentes S, Wood TA, Lawrenson T, et al. Gibberellins control fruit patterning in *Arabidopsis thaliana*. *Genes Dev.* 2010;24(19):2127–32. <https://doi.org/10.1101/gad.593410>.
66. Carrera E, Ruiz-Rivero O, Peres LEP, Atares A, Garcia-Martinez JL. Characterization of the procerato tomato mutant shows novel functions of the *SIDELLA* protein in the control of flower morphology, cell division and expansion, and the auxin-signaling pathway during fruit-set and development. *Plant Physiol.* 2012;160(3):1581–96. <https://doi.org/10.1104/pp.112.204552>.
67. Dorcey E, Urbez C, Blázquez MA, Carbonell J, Perez-Amador MA. Fertilization-dependent auxin response in ovules triggers fruit development through the modulation of gibberellin metabolism in *Arabidopsis*. *Plant J.* 2009;58(2):318–32. <https://doi.org/10.1111/j.1365-3113.2008.03781.x>.
68. Martí C, Orzáez D, Ellul P, Moreno V, Granel A. Silencing of *DELLA* induces facultative parthenocarpy in tomato fruits. *Plant J.* 2008;52(5):865–76. <https://doi.org/10.1111/j.1365-3113.2007.03282.x>.
69. Fu F, Mao W, Shi K, Zhou Y, Tadao A, Yu J. A role of brassinosteroids in early fruit development in cucumber. *J Exp Bot.* 2008;59(9):2299–308. <https://doi.org/10.1093/jxb/ern093>.
70. Zhu J, Sae-Seaw J, Wang Z. Brassinosteroid signalling. *Development.* 2013;140(8):1615–20. <https://doi.org/10.1242/dev.060590>.
71. Schwabe WW, Mills JJ. Hormones and parthenocarpic fruit-set. *Hortic Abstr.* 1981.
72. Figueiredo DD, Batista RA, Roszak PJ, Khler C. Auxin production couples endosperm development to fertilization. *Nat Plants.* 2015;1(12):15184. <https://doi.org/10.1038/nplants.2015.184>.
73. Figueiredo DD, Khler C. Auxin: a molecular trigger of seed development. *Genes Dev.* 2018;32(7-8):479–90. <https://doi.org/10.1101/gad.312546.118>.
74. Habib S, Waseem M, Li N, Yang L, Li Z. Overexpression of *SIGRAS7* affects multiple behaviors leading to confer abiotic stresses tolerance and impacts gibberellin and auxin signaling in tomato. *Int J Genomics.* 2019;4051981:1–16. <https://doi.org/10.1155/2019/4051981>.
75. Liu Y, Huang W, Xian Z, Hu N, Lin D, Ren H, et al. Overexpression of *SIGRAS40* in tomato enhances tolerance to abiotic stresses and influences auxin and gibberellin signaling. *Front Plant Sci.* 2017;8:1659. <https://doi.org/10.3389/fpls.2017.01659>.
76. Sánchez C, Vielba JM, Ferro E, Covelo G, Solé A, Abarca D, et al. Two *SCARECROW-LIKE* genes are induced in response to exogenous auxin in rooting-competent cuttings of distantly related forest species. *Tree Physiol.* 2007;27(10):1459–70. <https://doi.org/10.1093/treephys/27.10.1459>.
77. Della RF, Fattorini L, D'Angeli S, Velocchia A, Del DS, Cai G, et al. *Arabidopsis SHR* and *SCR* transcription factors and *AUX1* auxin influx carrier control the switch between adventitious rooting and xylogenesis in plants and in vitro cultured thin cell layers. *Altamura MM Ann Bot.* 2015;115(4):617–28. <https://doi.org/10.1093/aob/mcu258>.
78. Wiśniewska A, Pietraszewska-Bogiel A, Zuzga S, Tagashira N, Łotocka B, Malepszy S, et al. Molecular characterization of *SCARECROW (CSCR)* gene expressed during somatic embryo development and in root of cucumber (*Cucumis sativus* L.). *Acta Physiol Plant.* 2013;35:1483–95.
79. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools – an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant.* 2020;13(8):1194–202. <https://doi.org/10.1016/j.molp.2020.06.009>.
80. Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, et al. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res.* 2020;48(D1):D2658–D2668. <https://doi.org/10.1093/nar/gkz991>.
81. Punta M, Coghill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, et al. The Pfam protein families database. *Nucleic Acids Res.* 2012;40(D1):D290–301. <https://doi.org/10.1093/nar/gkr1065>.
82. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the ExPASy server. In: Walker JM, editor. *The proteomics protocols handbook*. Totowa: Humana Press; 2005. p. 571–607. <https://doi.org/10.1385/1-59259-890-0.571>.
83. Pérez-Rodríguez P, Riaño-Pachón DM, Corrêa LGG, Rensing SA, Kersten B, Mueller-Roeber B. PlnTFDB: updated content and new features of the plant transcription factor database. *Nucleic Acids Res.* 2010;38(suppl\_1):D822–D7. <https://doi.org/10.1093/nar/gkp805>.
84. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl Acids Res.* 2004;32(5):1792–7. <https://doi.org/10.1093/nar/gkh340>.
85. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 2016;33(7):1870–4. <https://doi.org/10.1093/molbev/msw054>.
86. Marchler-Bauer A, Bryant SH. CD-search: protein domain annotations on the fly. *Nucleic Acids Res.* 2004;32(Web Server):W327–W31. <https://doi.org/10.1093/nar/gkh454>.
87. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, Deweese-Scott C, et al. CDD: a conserved domain database for the functional

- annotation of proteins. *Nucleic Acids Res.* 2011;39(Database):D225–D9. <https://doi.org/10.1093/nar/gkq1189>.
88. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. Meme suite: tools for motif discovery and searching. *Nucleic Acids Res.* 2009; 37(Web Server):W202–W8. <https://doi.org/10.1093/nar/gkp335>.
  89. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, et al. MScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 2012;40(7):e49.
  90. Ma W, Chen C, Liu Y, Zeng M, Meyers BC, Li J, et al. Coupling of microRNA-directed phased small interfering RNA generation from long noncoding genes with alternative splicing and alternative polyadenylation in small RNA-mediated gene silencing. *New Phytol.* 2017;217(4):1535–50. <https://doi.org/10.1111/nph.14934>.
  91. Chen C, Li J, Feng J, Liu B, Feng L, Yu X, et al. sRNAanno - a database repository of uniformly-annotated small RNAs in plants. *Hortic Res.* 2019; 8(1):45. <https://doi.org/10.1038/s41438-021-00480-8>.
  92. Kalvari I, Argasinska J, Quinones-Olvera N, Nawrocki EP, Rivas E, Eddy SR, et al. Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Res.* 2018;46(D1):D335–D42. <https://doi.org/10.1093/nar/gkx1038>.
  93. Kalvari I, Nawrocki EP, Argasinska J, Quinones-Olvera N, Finn RD, Bateman A, et al. Non-coding RNA analysis using the Rfam database. *Curr Protoc Bioinformatics.* 2018;62(1):e51. <https://doi.org/10.1002/cpbi.51>.
  94. Mano S, Nakamura T, Kondo M, Miwa T, Nishikawa S-i, Mimura T, et al. The Plant Organelles Database 3 (PODB3) update 2014: integrating electron micrographs and new options for plant organelle research. *Plant Cell Physiol.* 2014;55(1):e1.
  95. Langmead B. Aligning short sequencing reads with bowtie. *Curr Protoc Bioinformatics.* 2010;32:11.17.11–4.
  96. Brousse C, Liu Q, Beauclair L, Deremetz A, Axtell MJ, Bouché N. A non-canonical plant microRNA target site. *Nucleic Acids Res.* 2014;42(8):5270–9. <https://doi.org/10.1093/nar/gku157>.
  97. Wan CY, Wilkins TA. A modified hot borate method significantly enhances the yield of high-quality RNA from cotton (*Gossypium hirsutum* L.). *Anal Biochem.* 1994;223(1):7–12. <https://doi.org/10.1006/abio.1994.1538>.
  98. Zhong H, Chen J, Li C, Chen L, Wu J, Chen J, et al. Selection of reliable reference genes for expression studies by reverse transcription quantitative real-time PCR in litchi under different experimental conditions. *Plant Cell Rep.* 2017;30(4):641–53. <https://doi.org/10.1007/s00299-010-0992-8>.
  99. Livak K, Schmittgen T. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta Ct}$  method. *Methods.* 2000;25(4):402–8. <https://doi.org/10.1006/meth.2001.1262>.
  100. Bolger AM, Marc L, Bjoern U. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114–20. <https://doi.org/10.1093/bioinformatics/btu170>.
  101. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29(1):15–21. <https://doi.org/10.1093/bioinformatics/bts635>.
  102. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotech.* 2015;33(3):290–5. <https://doi.org/10.1038/nbt.3122>.
  103. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550. <https://doi.org/10.1186/s13059-014-0550-8>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

