Evolutionary Comparison of Two Combinatorial Regulators of SBP-Box Genes, MiR156 and MiR529, in Plants

Shu-Dong Zhang1☯, Li-Zhen Ling2َا* , Quan-Fang Zhang3, Jian-Di Xu4, Le Cheng2َا*

1 Germplasm Bank of Wild Species, Kunming, 650201, China, 2 BGI-Yunnan, BGI-Shenzhen, Kunming, 650106, China, 3 Bio-Tech Research Center, Shandong Academy of Agricultural Sciences, Jinan 250100, China, 4 Shandong rice research institute, Shandong Academy of Agricultural Sciences, Jinan, 250100, China

☯ These authors contributed equally to this work.
* linglizhen@genomics.cn (L-ZL); chengle@genomics.cn (LC)

Abstract

A complete picture of the evolution of miRNA combinatorial regulation requires the synthesis of information on all miRNAs and their targets. MiR156 and miR529 are two combinatorial regulators of squamosa promoter binding protein-like (SBP-box) genes. Previous studies have clarified the evolutionary dynamics of their targets; however, there have been no reports on the evolutionary patterns of two miRNA regulators themselves to date. In this study, we investigated the evolutionary differences between these two miRNA families in extant land plants. Our work found that miR529 precursor, especially of its mature miRNA sequence, has a higher evolutionary rate. Such accelerating evolution of miR529 has significantly effects on its structural stability, and sequence conservation against existence of itself. By contrast, miR156 evolves more rapidly in loop region of the stable secondary structure, which may contribute to its functional diversity. Moreover, miR156 and miR529 genes have distinct rates of loss after identical duplication events. MiR529 genes have a higher average loss rate and asymmetric loss rate in duplicated gene pairs, indicating preferred miR529 gene losses become another predominant mode of inactivation, that are implicated in the contraction of this family. On the contrary, duplicated miR156 genes have a low loss rate, and could serve as another new source for functional diversity. Taken together, these results provide better insight into understanding the evolutionary divergence of miR156 and miR529 family in miRNA combinational regulation network.

Introduction

MicroRNAs (miRNAs) are small, non-coding RNA molecules that regulate gene expression by binding to target mRNA transcripts, leading to either translational repression or mRNA degradation. A growing body of evidence indicates that, in plants, a single miRNA can target and regulate multiple transcripts and conversely, the same genes can be targeted by a number of
Materials and Methods

Sequences of miR156 and miR529 in plants

Although over 300 miR156 sequences are registered in miRBase database (release 21) [11], only 43 high-confidence entries were collected for this study. It is reported that miR529 exists in several land plant organisms, but some miR529 genes had been obtained by similarity search and had not been validated with sufficiently experimental evidence prior to their addition to miRBase. To eliminate potential inaccuracies, these miR529 sequences were excluded from our analyses. Finally, a total of 11 miR529 members were selected from four plant species: Physcomitrella patens, Oryza sativa subsp. japonica, Brachypodium distachyon, and Zea mays. All precursor and mature sequences of miR156 and miR529 were downloaded from miRBase (release...
The information on miR156 and miR529 genes used in this study is summarized in S1 Table.

**Evaluation of substitution rate**

The precursor sequences of miR156 and miR529 were separately aligned using ClustalX and adjusted manually. The refined precursor alignments were used to calculate substitution rates using MEGA with the Kimura 2-parameter model. Previous studies indicated that considerable heterogeneity exists in the relative rates of evolution of different regions of the precursor miRNA sequences. Therefore, we parsed these precursor sequences into four regions as described by Shen: mature miRNA, miRNA complement (miRNA’), loop end, and stem extension (a stem structure, beyond the Dicer cut site). The substitution rate of each constituent part was calculated with the method described above. The independent samples t-test is used for statistical significance.

**Computation of folding energy and sequence identity of miR156 and miR529 precursors**

RNAfold was used to compute the minimal free energy (MFE) of each miR156 and miR529 precursor structure using default settings. Moreover, we eliminated the influence of the length of the miRNA precursor sequence by normalizing MFE as previously described. In this method, the normalized MFE (NMFE) was defined as the MFE divided by the sequence length of the miRNA precursor.

The sequence identity of precursor alignments of miR156 and miR529 was estimated via the online software Clustal Omega with the default setting. The same method was also applied to calculate sequence identity of mature alignments of miR156 and miR529.

**Detecting miR156 and miR529 expansions, dating duplication events, and estimating rates of gene loss**

Tandem duplication and segmental duplication significantly contribute to gene family expansion. Thus, we mainly focused on these two patterns of gene expansion in this study. Tandem duplications are characterized as multiple members of a gene family occurring within the same or neighboring intergenic regions and segmental duplications are defined as segments of DNA with near-identical sequence that map to two or more genomic locations. Of the four organisms analyzed in this study, we chose rice as the representation for analyzing miR156 and miR529 gene duplications because it has no less than two members in each miRNA family and a good genome annotation. Segmental duplications of miR156 and miR529 genes in rice genomes were retrieved from the Plant Genomic Duplication Database (PGDD) dataset. In this dataset, many protein-coding gene pairs excluding genes encoding miRNAs were cataloged in each duplicated block. In theory, if the members of a miRNA family reside within a duplicated block, then neighboring protein-coding genes would also be present on the same duplicated block. We therefore used 10 protein-coding genes both upstream and downstream of each miR156 and miR529 gene as guides to identify the duplicated block in which a miR156 or miR529 gene resides. The syntenic relationship between each miRNA gene pair in the duplicated block was visualized using MicroSyn.

The rates of synonymous substitution (Ks) of duplicated genes are expected to be similar over time. Therefore, we used Ks values to estimate the time at which the segmental duplication events took place. The Ks value of each gene pair within a duplicated block was extracted.
from the PDGG dataset [21]. The mean Ks value was calculated and used to date the duplication events. In this analysis, any Ks values greater than 2 were discarded because of the risk of saturation [16]. The approximate time at which the duplication event took place was then calculated using the mean Ks and an estimated rate of silent-site substitutions of $6.5 \times 10^{-9}$ substitutions/synonymous site/year [25].

The rate of gene loss was estimated according to Wang’s method [26]. The two copies in a duplicated block are referred to as copy 1 and copy 2 and the rate of gene loss in copy 2 was estimated as follows:

$$\frac{S_1}{S_1 + N_2} \times 100\%$$

Where, $S_1$ is the number of single-copy genes in copy 1; $N_2$ is the number of the extant genes in copy 2.

**Results**

**A higher evolutionary rate of miR529 in the mature sequence**

It has been reported that ubiquitously expressed genes evolve more slowly than tissue-specific genes, which suggests that the extent to which genes are expressed is critical for their evolutionary rates in multicellular organisms [27]. Therefore, we conjectured that narrowly expressed miR529 genes might have higher evolutionary rate than broadly expressed miR156 genes. To test this hypothesis, we first measured the evolutionary rates of the precursor miRNA (pre-miRNA) between the two miRNA families using MEGA with the Kimura 2-parameter model [14,15]. The result shows that the average evolutionary rate of miR529 precursor sequences was higher than that of miR156 precursor sequences; however the differences were not statistically significant (p > 0.05). Subsequently, we measured the independent evolutionary rates for the four parts of the precursor miRNA. In general, mature miRNA, as the functional part of the precursor miRNA, was well conserved between distant lineages and fewer mutations were found within its sequence, while the loop was the most variable region of the precursor sequence. As may be expected, the evolutionary rate was highest in the loop-containing region in both miR156 and miR529 families (Fig 1). Interestingly, the observed evolutionary rate in this region was significantly lower in the miR529 family members than in the miR156 family members (p < 0.01). The evolutionary rates in the stem and miRNA* regions were comparable between both miRNA families. Unexpectedly, the evolutionary rate of miR529 sequences was two times higher than that of miR156 sequences in the mature miRNA region (p < 0.01), which is contrary to our findings in the loop region. Collectively, these data suggest that miR529 has higher evolutionary rate in precursor, especially in mature sequences, whereas an accelerating evolution of miR156 precursor occurred in its loop region.

**Unstable secondary structure and weak conservation of miR529 sequences**

According to the above analyses, we knew that miR529 shows higher evolutionary rate than miR156 in its precursor sequence; however the difference is small (Fig 1). Therefore, it requires additional investigation to seek out reliable evidence. Structural studies revealed that the rapid evolution of miRNA precursors greatly affects the stability of secondary structure by introducing sequence variations. With this idea, we separately predicted the minimum free energy (MFE) of secondary structures for pre-miR529 and pre-miR156 using RNAfold [19]. Our thermodynamic data show that pre-miR529 sequences, on average, form more unstable secondary structures than pre-miR156 sequences, which generally have more negative MFEs (p < 0.01).
After normalizing the MFEs for sequence length, the normalized MFEs (NMFE) showed the same trend as described above (Fig 2B).

(Fig 2A). After normalizing the MFEs for sequence length, the normalized MFEs (NMFE) showed the same trend as described above (Fig 2B).

Fig 1. Comparison of evolutionary rates of *miR156* and *miR29* precursor sequences and their four structural elements. Error bars indicate the standard error of the mean.

doi:10.1371/journal.pone.0124621.g001

Fig 2. Energetic properties of the secondary structure of *miR156* and *miR529* precursors. Minimum free energy (MFE) (A) and Normalized minimum free energy (NMFE) (B). Error bars indicate the standard error of the mean.

doi:10.1371/journal.pone.0124621.g002
In addition, sequence conservation can be taken as a good indicator of evolutionary rate: the slower the sequence evolution, the higher the sequence conservation. The opposite situation is also true. For precursors, we pairwise compared full-length sequences of miR156 and miR529. As shown in S1A Fig, the precursor sequence identity of miR156 was higher than that of miR529 (p < 0.01). When only the mature sequences were considered, the difference of sequence identity between miR156 and miR529 was considerably dramatic. Pairwise sequence identity was over 95% and 77% for mature miR156 and miR529, respectively (S1B Fig). Therefore, these results provide additional evidence to support that miR156 family members were more evolutionarily constrained with a slow rate, whereas miR529 family members evolved more rapidly, particularly in the mature region of the precursor sequence.

Rapid rate of gene loss in miR529 after same duplication event in rice

Previous studies suggest that tandem and segmental duplications dominated the expansion of the miR156 family [28, 29]. As an ancient family, it is reasonable to hypothesize that miR529 has undergone a history of expansion events similar to miR156, which may be the underlying mechanism behind the amplification and diversification of this family. Therefore, we tested this hypothesis that duplication events played a role in the evolution of the miR529 family. As shown in Fig 3, ten of eleven miRNA genes within miR156 family formed duplicated pairs and were distributed in four duplicated blocks (blocks 19, 49, 57, and 162). This observation agreed with previous reports suggesting that the miR156 family mainly arose from large scale segmental duplication [29]. As for the miR529 family, two miR529 family members (miR529a and miR529b) reside within the duplicated block 50, indicating that miR529 genes also originated from segmental duplication. In addition, two duplicated gene pairs (miR156b/miR156c and miR156h/miR156j) of the miR156 family are located within the same gene and have been characterized as tandem duplications. However, the tandem duplication cannot be inferred because the two miR529 members in rice were characterized as segmental duplicated pairs. As a result, we investigated whether tandem duplication was the mechanism by which multiple members of the miR529 family arose in moss, though its genome is not well annotated. Our results indicate that three miR529 gene pairs of seven genes were located in neighboring regions and the distance between each miRNA pair was less than 100 nt, which satisfies the criteria for tandem duplication (data not shown). Consequently, we can infer that similar duplication mechanisms, including segmental duplication and tandem duplication are at work in the evolution of the miR529 family, as is the case with the miR156 family.

In addition, we used synonymous silent substitutions per site (Ks) as a proxy for time to estimate the approximate time of occurrence of the segmental duplication events. As can be seen from Table 1, the mean Ks values of the duplicated genes in block 50 in which miR529 resides are almost at the same level as the Ks values calculated for the other four blocks. This means that the segmental duplication events involving miR156 and miR529 genes occurred approximately 70 million years (Myr) ago, which is consistent with the time at which genome duplication events took place in rice [26, 30]. Therefore, these results suggested that miR156 and miR529 genes produced in the same genome duplication.

However, some genes in the duplicated blocks were frequently lost after the genome-wide duplication. For example, miR156b and miR156c are tandem duplications, yet there is only one corresponding miRNA, miR156l, residing within duplicated block 19. To help estimate the loss rate of the miR156 and miR529 genes, we tabulated the number of flanking protein-coding genes of one miRNA copy that does not have its counterparts in the other duplicated copy within the same block. Then, the rates of gene loss were compared between blocks and between two copies of the same block. As seen in Table 2, the block 50 where miR529 genes resided had a higher
average rate of loss than other four duplicated blocks (blocks 19, 49, 57, and 162) in which miR156 genes are located (Table 2). In addition, we found that two copies of miR529 duplicated pairs have greatly asymmetric rates of gene loss. One copy (copy 2) showed the highest percentage (70%) of gene loss, while the other copy had only 28% of gene loss. By contrast, the majority of miR156 duplicated gene pairs showed very little difference of the gene loss rate between the two copies (Table 2). All together, our results indicate that two miR529 and miR156 families experienced the same genome-wide duplication event, and yet exhibited the different rate of gene loss. The rapidly evolving miR529 family had more asymmetric rate of gene loss between the two copies with a higher average rate of loss than miR156 family.

Fig 3. Syntenic duplicated paralogs of miR156 and miR529 genes in the duplicated blocks in rice. Black arrowheads indicate the positions of duplicated miRNA pairs in every duplicated block. The flanking protein-coding gene pairs are linked by grey lines.

doi:10.1371/journal.pone.0124621.g003

Table 1. Identified duplicated blocks containing miR156 and miR529 and estimation of the absolute date for segmental duplication events in rice.

| Duplicated pair       | Block | Mean Ks | SD Ks | Date (Myr) |
|-----------------------|-------|---------|-------|------------|
| osa-mir156bc/l        | 19    | 0.798   | 0.053 | 66.159     |
| osa-mir156e/i         | 49    | 0.764   | 0.031 | 63.368     |
| osa-mir529a/b         | 50    | 0.863   | 0.026 | 71.570     |
| osa-mir156d/hj        | 57    | 0.821   | 0.032 | 68.102     |
| osa-mir156f/g         | 162   | 0.784   | 0.108 | 65.030     |

doi:10.1371/journal.pone.0124621.t001
Discussion

How do complex regulatory networks evolve and how does their evolution result in phenotypic change and speciation? For a long time, evolutionary biologists have been devoted to solving these questions. A good approach to understanding the evolution of the gene regulation network is to synthesize the information about all individual regulators and their targets. In our previous study, the SBP-box genes targeted by miR529 have been found to evolve differently from those targeted by miR156 in plants [6]. However, the evolutionary patterns of their two combinational regulators, miR156 and miR529 are still unknown in plants. Growing evidence has revealed that miR156 and miR529 exhibit significant differences in gene expression patterns, taxonomic distribution, and the number of members in the miRNA family (See details in Introduction section). These different characteristics between miR156 and miR529 drove us to explore whether they evolve in different patterns in plants.

Our results suggest that narrowly expressed miR529 genes have higher evolutionary rates in precursor sequences than broadly expressed miR156 genes. The difference in evolutionary rate between them was small, but additional evidence found by analyzing the structural stability and sequence conservation support this small difference (Fig 1). Therefore, the expression patterns of miR156 and miR529 genes were negatively correlated with their respective rates of sequence evolution. This negative correlation between expression patterns and evolutionary rates has also been found in Long Intergenic Noncoding RNAs (LncRNAs) [31], indicative of a general feature of gene evolution. In parallel, the difference of their expression patterns in specific tissues and developmental stages facilitated the combinatorial regulation of their common target genes during evolution.

Our results also revealed that mature miRNA and loop sequences in four miRNA precursor elements change at different rates in the two analyzed miRNA families. The loop sequences of miR529 have a lower evolutionary rate than those of miR156 and the opposite trend was observed in the mature miRNA sequences (Fig 1). A large body of evidence has proven that variations in the loop sequences of miRNA precursors can contribute to phenotypic variations by altering miRNA regulation. Wang et al. [32] found that a GG/AA polymorphism in the loop structure of miR2923a was correlated to the seed length of two cultivated rice varieties: japonica and indica. Another study revealed that the alterations in pre-miR181 loop sequences can modulate the activities of miR181 family members, albeit of nearly identical or identical mature miRNAs [2]. Moreover, the studies illustrated that the loop region can affect miRNA gene levels by influencing miRNA processing [33]. Accordingly, the question of whether the more variable miR156 loop sequences may have changed the regulatory activity of its family members by impairing miRNA processing is worthy of additional investigation. On the other hand, variations in mature miRNA can disrupt base pairing to the miRNA* and target sequences. First, the mismatches between miRNA and miRNA* introduced by variations in the mature miRNA sequence can destabilize the structure of the miRNA precursor. This has been confirmed by our analysis of the thermodynamic stability of miRNA precursors (Fig 2A and 2B). Similarly,

| Family | Block | Copy 1 | Copy 2 | Average |
|--------|-------|-------|-------|---------|
| miR156 | 19    | 0.488 | 0.416 | 0.452   |
| miR156 | 49    | 0.414 | 0.479 | 0.446   |
| miR529 | 50    | 0.283 | 0.701 | 0.492   |
| miR156 | 57    | 0.337 | 0.585 | 0.461   |
| miR156 | 162   | 0.407 | 0.505 | 0.456   |

doi:10.1371/journal.pone.0124621.t002
variations in the mature sequence also disrupt binding to the target sequence. In our previous study, miR529 targets are more conserved and evolve at a slower rate than miR156 targets [6]. In this case, frequent variations only in mature miR529 sequences may lead to the loss of their target regulation but not miR156 target regulation. Additionally, studies have demonstrated that the variations in mature miRNA sequences can abolish the production of mature miRNAs by disrupting miRNA processing [34,35]. Thus, the subsequent loss of miR529 regulation may occur through mutations in the mature miRNA during the evolution of core eudicots. Nevertheless, whether or not miR529 genes block their own biogenesis by altering the secondary structure of pre-miRNAs is still unresolved and requires further studies.

Gene duplication is a major route of origination for miRNAs. Without exception, two ancient miR156 and miR529 families were expanded through the same segmental duplication as well as tandem duplication. However, our data also illustrate that miR156 and miR529 genes have significantly different rates of gene loss after duplication. Besides a higher average loss rate, miR529 genes had an obviously asymmetric rate of gene loss in two duplicated copies (Table 2). This asymmetric acceleration of the evolutionary rate can bring about more mutations in one of the paralogs and lead to its more rapid loss [36–38]. Therefore, the high and asymmetric gene loss rate is a predominant mode of inactivation of duplicated miR529 genes and contributes to reducing the impact of a genome-wide duplication event on functional redundancy within miR529 family members. On the contrary, the majority of miR156 family members with low loss rate could be readily retained after duplication, which contributed to the expansion of this family. Moreover, it is proven that these retained miR156 family members have functionally diverged and exhibit diverse expression profiles in Arabidopsis [28]. Accordingly, the gene duplication and subsequent divergence of miR156 family members could serve as new sources for functional diversity and confer phenotypic differentiation in development.

Altogether, a combination of the rapid evolution of miR529 sequence (especially the mature sequence), and the asymmetric rate of gene loss makes up the important evolutionary forces that cause the contraction of this miRNA family. An extreme phenomenon is that all the miR529 members are extinct in core eudicots. Our previous work demonstrated stronger purifying selection against mutations within the binding sites of miR529 targets [6]. Meanwhile, these miR529 targets also are cooperatively controlled by miR156. If the mutations occurred on the binding site of miR529 targets, they would be bound to disrupt ligand binding between miR156 and this same target. Therefore, the economical and balanced way to change miR529 combinational regulation is to change miR529 itself. Indeed, the retention of miR529 binding sites in some members of SBP-box genes but no miR529 candidate in core eudicots gave it a full interpretation (S2 Table). On the contrary, the increased mutations in miR156 targets but the relative conservation in miR156 genes should contribute to harmonizing the miR156 regulation system. For miR156, the variable loop sequence followed by the divergence of duplicated gene pairs after duplication might be the principle manner in which miR156 functions are enriched. Moreover, accumulating functional studies have revealed that the SBP-box genes targeted by miR156 and miR529 are involved in various development processes and lead to many morphological differences, such as leaf, flower, pollen, and fruit (see review [39]). It is well known that these important characteristics might distinguish monocot and dicot from each other. Therefore, the differences between miR156 and miR529 regulation mechanisms might contribute to the morphological divergence of monocot and dicot. In conclusion, our data indicate that miR156 and miR529 genes show different evolutionary dynamics. MiR156 genes are continually differentiated and amplified by accelerated mutations in their hairpin loops and diverging duplicated gene pairs after duplication. By contrast, miR529 genes gradually reduce their abundance via accumulating more mutations in the mature miRNA sequences and by rapidly losing one of the paralogous duplicated pairs soon after the duplication event. Taken
together, these results enhance our understanding of the different evolutionary mechanisms driving changes in two miRNA families and the different regulatory mechanisms in the gene regulation network.

Supporting Information

S1 Fig. Comparisons of sequence identity between miR156 and miR529 genes in precursor (A) and mature (B) sequences. Error bars indicate the standard error of the mean.
(TIF)

S1 Table. High-confidence entries of miR156 and miR529 family used in this study.
(DOC)

S2 Table. Predicted targets for miR529 in three eudicots species.
(DOC)

Acknowledgments

We thank Jinyong Hu for constructive suggestions and Yunlong Liu for discussion of data analysis.

Author Contributions

Conceived and designed the experiments: L-ZL LC. Performed the experiments: S-DZ Q-FZ J-DX LC. Analyzed the data: S-DZ L-ZL. Contributed reagents/materials/analysis tools: S-DZ L-ZL. Wrote the paper: S-DZ L-ZL.

References

1. Jeong DH, Park S, Zhai J, Gurazada SG, De Paoli E, Meyers BC, et al. (2011) Massive analysis of rice small RNAs: mechanistic implications of regulated microRNAs and variants for differential target RNA cleavage. Plant Cell 23: 4185–4207. doi:10.1105/tpc.111.089045 PMID: 22158467

2. Liu G, Min H, Yue S, Chen CZ (2008) Pre-miRNA loop nucleotides control the distinct activities of mir-181a-1 and mir-181c in early T cell development. PLoS One 3: e3592. doi:10.1371/journal.pone.0003592 PMID: 18974849

3. Zheng Y, Li YF, Sunkar R, Zhang W (2012) SeqTar: an effective method for identifying microRNA guided cleavage sites from degradome of polyadenylated transcripts in plants. Nucleic Acids Res 40: e28. doi:10.1093/nar/gkr1092 PMID: 22140118

4. Chuck G, Whipple C, Jackson D, Hake S (2010) The maize SBP-box transcription factor encoded by tasselsheath4 regulates bract development and the establishment of meristem boundaries. Development 137: 1243–1250. doi:10.1242/dev.048348 PMID: 20223762

5. Cuperus JT, Fahlgren N, Carrington JC (2011) Evolution and functional diversification of MiRNA genes. Plant Cell 23: 431–442. doi:10.1105/tpc.110.082784 PMID: 21317375

6. Ling LZ, Zhang SD (2012) Exploring the evolutionary differences of SBP-box genes targeted by miR156 and miR529 in plants. Genetica 140: 317–324. doi:10.1007/s10709-012-9684-3 PMID: 23054224

7. Jeong DH, Schmidt SA, Rymarquis LA, Park S, Ganssmann M, German MA, et al. (2013) Parallel analysis of RNA ends enhances global investigation of microRNAs and target RNAs of Brachypodium distachyon. Genome Biol 14: R145. doi:10.1186/gb-2013-14-12-r145 PMID: 24367943

8. Zhang L, Chia JM, Kumari S, Stein JC, Liu Z, Narechania A, et al. (2009) A genome-wide characterization of microRNA genes in maize. PloS Genet 5: e1000716. doi:10.1371/journal.pgen.1000716 PMID: 19936050

9. Arazi T, Talmor-Neiman M, Stav R, Riese M, Huijser P, Baulcombe DC (2005) Cloning and characterization of micro-RNAs from moss. Plant J 43: 837–848. PMID: 16146523

10. Axtell MJ, Snyder JA, Bartel DP (2007) Common functions for diverse small RNAs of land plants. Plant Cell 19: 1750–1769. PMID: 17601824
11. Kozomara A, Griffiths-Jones S (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res 42: D68–73. doi: 10.1093/nar/gkt1181 PMID: 24275495

12. Barakat A, Wall K, Leebens-Mack J, Wang YJ, Carlson JE, Depamphilis CW (2007) Large-scale identification of microRNAs from a basal eudicot (Eschscholzia californica) and conservation in flowering plants. Plant J 51: 991–1003. PMID: 17635767

13. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25: 4876–4882. PMID: 9396791

14. Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16: 111–120. PMID: 7463489

15. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30: 2725–2729. doi: 10.1093/molbev/ms317 PMID: 24132122

16. Guo X, Su B, Zhou Z, Sha J (2009) Rapid evolution of mammalian X-linked testis microRNAs. BMC Genomics 10: 97. doi: 10.1186/1471-2164-10-97 PMID: 19257908

17. Mohammed J, Flynt AS, Siepel A, Lai EC (2013) The impact of age, biogenesis, and genomic clustering on Drosophila microRNA evolution. RNA 19: 1295–1308. doi: 10.1261/rna.039248.113 PMID: 23882112

18. Shen Y, Lv Y, Huang L, Liu W, Wen M, Tang T, et al. (2011) Testing hypotheses on the rate of molecular evolution in relation to gene expression using microRNAs. Proc Natl Acad Sci U S A 108: 15942–15947. doi: 10.1073/pnas.1109801108 PMID: 21911382

19. Lorenz R, Bernhart SH, Honer Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, et al. (2011) ViennaRNA Package 2.0. Algorithms Mol Biol 6: 26. doi: 10.1186/1748-7188-6-26 PMID: 22115189

20. Zhu Y, Skogerbo G, Ning QQ, Wang Z, Li BQ, Yang S, et al. (2012) Evolutionary relationships between miRNA genes and their activity. BMC Genomics 13.

21. Lee TH, Tang H, Wang X, Paterson AH (2013) PGDD: a database of gene and genome duplication in plants. Nucleic Acids Res 41: D1152–1158. doi: 10.1093/nar/gks1104 PMID: 23180799

22. Vision TJ, Brown DG, Tanksley SD (2000) The origins of genomic duplications in Arabidopsis. Science 290: 2116–2117. PMID: 1118139

23. Cai B, Yang X, Tuskan GA, Cheng ZM (2011) MicroSyn: a user friendly tool for detection of microsynteny in a gene family. BMC Bioinformatics 12: 79. doi: 10.1186/1471-2105-12-79 PMID: 21418570

24. Shi SH, Karlowsk WM, Pan R, Tzen YH, Mayer KF, Li WH (2004) Comparative analysis of the receptor-like kinase family in Arabidopsis and rice. Plant Cell 16: 1220–1234. PMID: 15105442

25. Yu J, Wang J, Lin W, Li S, Li H, Zhou J, et al. (2005) The Genomes of Oryza sativa: a history of duplications. PloS Biol 3: e38. PMID: 15685292

26. Wang X, Shi X, Hao B, Ge S, Luo J (2005) Duplication and DNA segmental loss in the rice genome: implications for diploidization. New Phytol 165: 937–946. PMID: 15720704

27. Yang J, Su AI, Li WH (2005) Gene expression evolves faster in narrowly than in broadly expressed mammalian genes. Mol Biol Evol 22: 2113–2118. PMID: 15987875

28. Maher C, Stein L, Ware D (2006) Evolution of Arabidopsis microRNA families through duplication events. Genome Res 16: 510–519. PMID: 16520461

29. Wang S, Zhu OH, Guo X, Gui Y, Bao J, Helliwell C, et al. (2007) Molecular evolution and selection of a gene encoding two tandem microRNAs in rice. FEBS Lett 581: 4789–4793. PMID: 17884044

30. Paterson AH, Bowers JE, Chapman BA (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc Natl Acad Sci U S A 101: 9903–9908. PMID: 15161969

31. Managadze D, Rogozin IB, Chernikova D, Shabalina SA, Koonin EV (2011) Negative correlation between expression level and evolutionary rate of long intergenic noncoding RNAs. Genome Biol Evol 3: 1390–1404. doi: 10.1093/gbe/evr116 PMID: 22071789

32. Zhang X, Zeng Y (2010) The terminal loop region controls microRNA processing by Drosha and Dicer. Nucleic Acids Res 38: 7689–7697. doi: 10.1093/nar/gkq645 PMID: 20660014

33. Duan R, Pak C, Jin P (2007) Single nucleotide polymorphism associated with mature miR-125a alters the processing of pri-miRNA. Hum Mol Genet 16: 1124–1131. PMID: 17400653

34. Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, et al. (2003) The nuclear Rhase III Drosha initiates microRNA processing. Nature 425: 415–419. PMID: 14508493
36. Brunet FG, Roest Crollius H, Paris M, Aury JM, Gibert P, Jaillon O, et al. (2006) Gene loss and evolutionary rates following whole-genome duplication in teleost fishes. Mol Biol Evol 23: 1808–1816. PMID: 16809621

37. Byrne KP, Wolfe KH (2007) Consistent patterns of rate asymmetry and gene loss indicate widespread neofunctionalization of yeast genes after whole-genome duplication. Genetics 175: 1341–1350. PMID: 17194778

38. Yampolsky LY, Bouzinier MA (2014) Faster evolving Drosophila paralogs lose expression rate and ubiquity and accumulate more non-synonymous SNPs. Biol Direct 9: 2. doi: 10.1186/1745-6150-9-2 PMID: 24438455

39. Preston JC, Hileman LC (2013) Functional Evolution in the Plant SQUAMOSA-PROMOTER BINDING PROTEIN-LIKE (SPL) Gene Family. Front Plant Sci 4: 80. doi: 10.3389/fpls.2013.00080 PMID: 23577017