

## CONVERGENCE AND DIVERGENCE STUDIES OF PLANT PRECURSOR MicroRNAs

ABDUL GHANI, MUHAMMAD DIN\* AND MUHAMMAD YOUNAS KHAN BAROZAI

*Department of Botany, University of Balochistan, Sariab Road Quetta, Pakistan*

*\*Corresponding author's email: drmdin75@gmail.com*

### Abstract

MicroRNA (miRNA) is a class of small RNAs. It is non-coding, 18-26 nucleotides (nt) in length and play a useful and vital role in post-transcriptional gene regulation by targeting messenger RNAs (mRNAs). The short mature miRNAs generate from long precursor miRNAs (pre-miRNAs). The length of pre-miRNAs is range from 50-500 nt. Many of the mature miRNAs are evolutionarily conserved but the convergence of the pre-miRNAs is not well studied. The aim of this research is to study the convergence and divergence of pre-miRNAs in the significant and major groups of plants. In this research, homology-based comparative genomics approach was employed to study the convergence and divergence studies of pre-miRNAs from 35 species of plants. Among the selected species, two are from bryophyta and pteridophyta, one from each, three from gymnosperm, three from monocot, and twenty seven species were from dicot. The three monocots, containing 969 pre-miRNAs, were subjected to convergence and divergence analyses. Out of 969, 27 pre-miRNAs showed convergence with gymnosperm, 26 with bryophyta and pteridophyta and 157 with dicot. Similarly, out of 104, 20 pre-miRNAs of gymnosperm showed convergence with monocot, eight with bryophyta and pteridophyta, and 26 with dicot. The 27 out of 287 pre-miRNAs of bryophyta and pteridophyta showed convergence with monocot, five with gymnosperm, and 18 with dicot. The 148 out of 2647 pre-miRNAs of dicot showed convergence with monocot, 62 with gymnosperm and 51 with bryophyta and pteridophyta. These findings would help us to better understand the pre-miRNAs convergence and divergence in plants. As pre-miRNA generate the functional mature miRNA, so, findings of pre-miRNAs convergence and divergence will also be useful to design and regulate the miRNAs expression for better crop production, biotic and abiotic stress management at the pre-miRNAs level in plants.

**Key words:** Bryophyta, Comparative genomics, Gymnosperm, Pre-miRNAs, Pteridophyta.

### Introduction

MicroRNAs (miRNAs) are endogenous, small in length and do not code for protein. They play an important role in gene regulation for many species of plants and animals. Since the first miRNA discovery in *Caenorhabditis elegans* (Lee *et al.*, 1993), detailed studies into transcription and the functional role of miRNAs across different species have led to a complex picture of miRNA biogenesis and miRNA-associated regulatory pathways (Bartel, 2004).

The miRNA biogenesis is a multistep process. Initially, miRNA gene is transcribed as a primary miRNA (pri-miRNA) by the help of RNA polymerase II enzyme. This pri-miRNA give rises into a precursor miRNA (pre-miRNA) by Dicer-like1 protein (DCL1) (Jones-Rhoades *et al.*, 2006). The Pre-miRNA generate a self-folded stem-loop structure of various length and levels of complexity (bulges, multiple shorter sub-hairpins etc.). Later, the pre-miRNA is processed also by Dicer-like1 proteins to produce 18-26 nt long double-stranded RNA duplex (miRNA:miRNA\*) and exported to the cytoplasm via an exportin homologue. In the cytoplasm, the RNA duplex denatures and the mature miRNA strand is incorporated into the RNA induced silencing complex (RISC) (Baloch *et al.*, 2013). The miRNA-RISC complex in association with members of the Argonaute (Ago) protein family target relevant mRNAs using base complementarity as a search tool. The miRNA pairing with the mRNA causes either mRNA degradation or suppression from translation and it is dependent on the level of complementarity between miRNA and its targeted mRNA (Voinnet, 2009).

Many studies suggest that pre-miRNAs often show diverse features regarding secondary structures, such as multiple miRNA/miRNA\* duplexes per precursor, multiple hairpin loops, and tRNA precursor-like clover structures (Babiarz *et al.*, 2008). The pre-miRNA sequence has great influence on the generation of the mature miRNA and ultimately on the gene regulation at the post-transcriptional stage. This reveals that pre-miRNAs play an important role in developing a better understanding of mechanisms related to miRNA origin and evolution (Jones-Rhoades *et al.*, 2006).

As mature miRNAs are highly conserved among the plant species and well-studied phenomenon in small RNA world. However, the nature of the pre-miRNAs convergence and divergence is not clear in the plant kingdom. This study is aimed to study the significant plant species pre-miRNAs convergence and divergence based on their sequences, available in the miRNA-database (miRNA registry database). A total of 4007 pre-miRNAs sequences were downloaded from miRNA registry database (Release 21: June 2014) (Griffiths-Jones, 2004) and subjected to convergence and divergence analyses based on the comparative genomics. These pre-miRNAs belonged to the 35 species of significant groups of plants like; dicot, monocot, gymnosperm, bryophyta and pteridophyta.

### Materials and Methods

**Fetching of pre-miRNAs sequences:** A sum of 4007 pre-miRNA sequences belonging to the 35 plant species was downloaded from the miRNA registry database (Release 21: June 2014) (Griffiths-Jones, 2004). The pre-miRNA

sequences with mature sequences were saved and further utilized as subject for homology based comparative genomics search analyses with similar approaches used earlier (Barozai 2012a, b, c, d; Barozai, 2013 ).

**Homology search by BLAST (Basic Local Alignment Search Tool):** In this step, a similar approach, as reported previously by Zhang *et al.*, (2006) and with modification by Barozai *et al.*, (2008), is applied. All the selected known pre-miRNAs of major plant groups were aligned with each other using Blastn program (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) available at National Center for Biotechnology Information (NCBI) (Altschul *et al.*, 1990), as shown in Fig. 1. The Blastn parameter settings were adjusted as follows: expect values ~1,000; low density was chosen as the sequence filter, database (others; nucleotide and organism; spp name), program selection (somewhat similar sequence) and all other parameters were used as default.

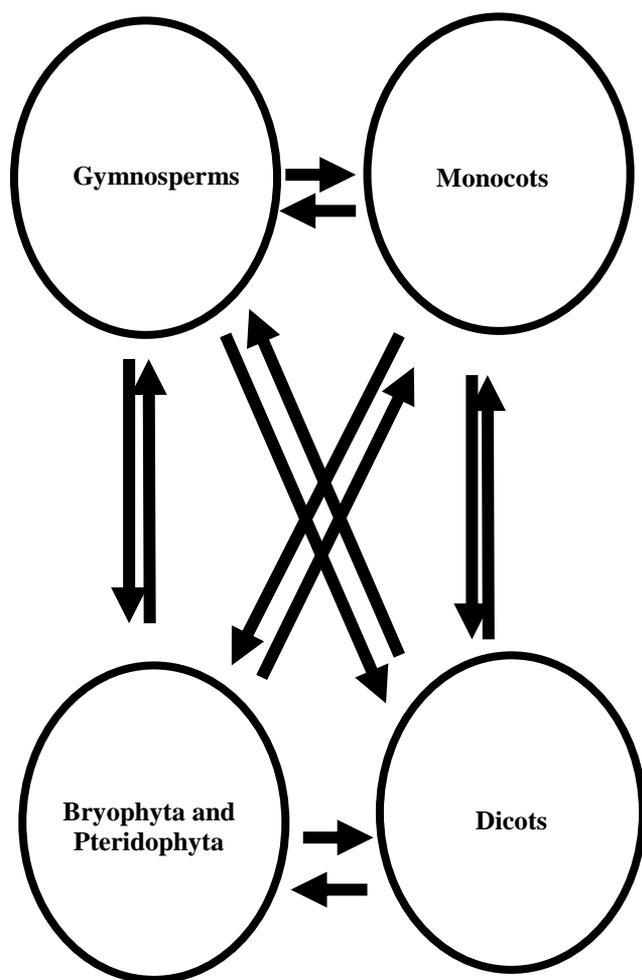


Fig. 1. Graphical representation showing the computational identification of convergence and divergence studies of plants precursor microRNAs among four groups of plants.

**Convergence and divergence studies:** The Blast software displayed the results in three forms such as graphical, tabular and alignment. The tabular form showed the results for maximum score, strand orientation, total score, query coverage, maximum identities, E-value and accession

number. During the homology search, the sequences having 0-4 mismatches for the mature miRNAs and whose query coverage is from 30-100% were considered for convergence analyses. While the pre-miRNA sequences having more than four mismatches in mature sequences and whose query coverage is less than 30% were considered as divergence.

## Results and Discussion

**Fetching of pre-miRNAs sequences:** The current research is a comparative genomics based approach using bioinformatics tools. This is in agreement with the other reports by many research groups (Barozai & Husnain, 2011; Barozai & Wahid, 2012; Barozai *et al.*, 2012a; Barozai *et al.*, 2014a, b; Barozai & Din, 2014; Behlil *et al.*, 2016; Barozai & Din, 2017; Jahan *et al.*, 2017; Shah *et al.*, 2017), where the similarity based search by utilizing comparative genomics is found a valid and rational method to report new findings in many organisms at genomics level. Here, a total of 49 novel miRNAs were identified. Here, a total of 4007 pre-miRNAs sequences were subjected for identifying convergence and divergence in plant pre-miRNAs. These sequences belonged to the 35 species of important groups of plants and their distribution is as; dicot (2647 pre-miRNAs from 27 species; *Cynara cardunculus* (48), *Helianthus annuus* (06), *Helianthus arqophyllus* (03), *Helianthus exilis* (02), *Helianthus paradoxus* (03), *Helianthus tuberosus* (16), *Helianthus ciliaris* (03), *Helianthus petiolaris* (03), *Arabidopsis thaliana* (325), *Arabidopsis lyrata* (205), *Brassica napus* (90), *Brassica rapa* (96), *Brassica oleracea* (10), *Acacia auriculiformis* (07), *Glycine soja* (13), *Theobroma cacao* (82), *Lotus japonicas* (62), *Gossypium herbaceum* (01), *Medicago truncatula* (672), *Phaseolus vulgaris* (08), *Acacia mangium* (03), *Gossypium arboretum* (01), *Vigna unguiculata* (18), *Gossypium hirsutum* (78), *Arachis hypogaea* (23), *Gossypium raimondii* (296) and *Glycine max* (573), monocot (969 pre-miRNAs from 3 spp; *Oryza sativa* (592), *Sorghum bicolor* (205) and *Zea mays* (172)), gymnosperm (104 pre-miRNAs from 3 spp; *Picea abies* (40), *Pinus densata* (29) and *Pinus taeda* (35)), bryophyta and pteridophyta (287 pre-miRNAs from 2 spp; *Physcomitrella patens* (229) and *Selaginella moellendorffii* (58)). All these known pre-miRNAs were downloaded from miRBase (Release 21: June 2014) (Griffiths-Jones, 2004).

**Homology based comparative genomics analyses:** Data mining, homology search and comparative genomics resulted in many putative and interesting findings in various plant species (Barozai *et al.*, 2011; Barozai *et al.*, 2012b; Bibi *et al.*, 2017; Gul *et al.*, 2017). In the current study, 2647 pre-miRNAs belonging to the 27 spp. of dicot plants were aligned with the nucleotide database of gymnosperm, monocot, bryophyta and pteridophyta at NCBI. Similarly, the 104 pre-miRNA sequences from the two spp. of gymnosperm were blast against nucleotide database of monocot, dicot, bryophyta and pteridophyta respectively. Later the selected 287 pre-miRNAs of two spp. of bryophyta and pteridophyta were subjected to convergence and divergence analyses against the nucleotide

database of gymnosperm, monocot and dicot.

**Table 1. The pre-miRNA sequences of the significant plant groups (Monocot, Gymnosperm, Dicot, Bryophyta and Pteridophyta).**

Plant groups	Monocot pre-miRNAs 969	Gymnosperm pre-miRNAs 104	Bryophyta and Pteridophyta pre-miRNAs 287	Dicot pre-miRNAs 2647
Monocot 969	-----	Converged=27 Diverged=879	Converged=26 Diverged=880	Converged=157 Diverged=749
Gymnosperm 104	Converged=20 Diverged=84	-----	Converged=8 Diverged=96	Converged=26 Diverged=78
Bryophyta and Pteridophyta 287	Converged=27 Diverged=260	Converged=5 Diverged=282	-----	Converged=18 Diverged=269
Dicot 2647	Converged=148 Diverged=2499	Converged=62 Diverged=2585	Converged=51 Diverged=2596	-----

Out of 2647 pre-miRNAs, the 51 pre-miRNAs of dicot were found converge with bryophyta and pteridophyta, 62 with gymnosperm and 148 with monocot. Whereas 26 pre-miRNAs (from 969 pre-miRNAs) of monocot were observed as converged with bryophyta and pteridophyta, 27 with gymnosperm and 157 with dicot. In gymnosperm, 8 pre-miRNAs (out of 104) were identified as converge with bryophyta and pteridophyta, 20 with monocot and 26 with dicot. Similarly, from 287 pre-miRNAs of bryophyta and pteridophyta, 18 pre-miRNAs were predicted as converge with dicot, 27 with monocot and 5 with gymnosperm as shown in (Table 1).

### Convergence and divergence studies

**Convergence and divergence of dicot with bryophyta and pteridophyta pre-miRNAs:** The twenty seven plants of the dicot containing 2647 pre-miRNAs were studied for convergence and divergence with the two plants of the bryophyta and pteridophyta; *P. patens*, *S. moellendorffi* having 287 pre-miRNAs in the miRBase (Release 21: June 2014). Only 51 out of 2647 pre-miRNAs were found as converged with bryophyta and pteridophyta. These pre-miRNAs are observed with 30% to 50% and 71% to 100% query coverage convergence. The maximum converged pre-miRNAs of dicot with bryophyta and pteridophyta were found with 30% to 50% query coverage. More convergence is observed between dicot and *S. moellendorffii* of pteridophyta. The converged pre-miRNAs are mir156a, 156b, 156c, 156d, 156v, 158a, 160a, 160b, 160c, 164d, 166a, 166b, 166d, 166i, 171a, 171d, 171f, 171g, 172a, 172c, 319f, 395, 395b, 396a, 396b, 396c, 396d, 399b, 400, 408, 862b, 1310, 1510a, 1514b, 2108b, 2111b, 4386, 4993, 4995, 5033, 5240, 5293, 5298a, 5370, 5670, 5783, 6032, 6034, 6110, 6299 and 7538. These findings of more pre-miRNAs convergence are reported for the first time but supported in patches by the various research reports (Barozai *et al.*, 2008; 2012b; Din & Barozai, 2014a, b; Din *et al.*, 2014; 2016). Similarly, the eleven miRNA families; miR156, 160, 166, 167, 171, 319, 390, 395, 408, 414 and 419 are reported as conserved between *A. thaliana*, *O. sativa* and *P. patens* (Baloch *et al.*, 2015a, b). The miRNA families; miR529 and 535 are conserved in *P. patens* and *O. sativa*. The thirteen miRNAs families belonging to *P. patens* (miR156, 160, 166, 167, 171, 319, 390, 395, 408, 477, 529, 535, 894) are conserved with one or more

angiosperms. Bryophytes are thought to have shared a common ancestor with flowering plants 400 million years ago and these miRNAs have played regulatory roles since that period. One family (ppt-miR536) is conserved with lycopods, and the 93 families are considered moss specific, since they have not yet been cloned from any other land plant (Arazi *et al.*, 2005).

Whereas, the majority of the reported dicot pre-miRNAs (2596 out of 2647) are observed as diverged with bryophyta and pteridophyta. This is an indication of the emergence of new miRNAs in plant kingdom with the course of evolution.

### Convergence and divergence between dicot and gymnosperms pre-miRNAs:

The 2647 pre-miRNAs of dicot were studied for convergence and divergence against three plants of gymnosperm. Only 62 pre-miRNAs are found as converged with pre-miRNAs of three species of gymnosperm. These pre-miRNAs are observed with 30% to 100% query coverage. The maximum converged pre-miRNAs between dicot and gymnosperm are observed with 30% to 50% query coverage. More convergence is observed between dicot and *P. tada*. The converged pre-miRNAs are; mir156a, 156b, 156e, 156j, 156k, 157a, 157b, 160b, 160d, 160f, 166a, 166e, 167d, 169b, 169c, 172b, 172c, 319a, 390, 390a, 390b, 390d, 394b, 395, 396e, 397a, 397b, 400, 415, 479, 859, 861, 1109, 1310, 1510a, 1535a, 1886, 2111a, 2111b, 2111o, 2630c, 2630y, 3436, 3476, 3510, 3520, 3630, 4995, 5020c, 5037b, 5204, 5233, 5288, 5298b, 5373, 5654, 5670, 5743a, 5743b, 5781, 5783 and 6032.

The miRNA 156 family member's sequences are reported converged in both gymnosperm and angiosperms (Yang *et al.*, 2007). Similarly, the miRNAs that are found in flowering plants like; miRNA 160 and 390, are also found in gymnosperm (Yang *et al.*, 2007). The miRNA 159 and 319 are involved in fertility and development. These both miRNA families are reported in angiosperms and gymnosperms (Luo *et al.*, 2013).

In this study, 2585 out of 2647 pre-miRNAs of the dicot, are observed as diverged between dicot and gymnosperm, showing pre-miRNAs based adaptation by dicot plants.

### Convergence and divergence between dicot and monocot pre-miRNAs:

To study convergence and divergence between dicot and monocot, 2647 pre-miRNAs of dicot were blast against the nucleotide database of three

species of monocot. 148 out of 2647 pre-miRNAs were found as converged pre-miRNAs with monocot. These pre-miRNAs are observed with 30% to 100% query coverage. The maximum converged pre-miRNAs between dicot and monocot are found with 30% to 50% query coverage. More convergence is observed between dicot and *O. sativa*. The converged pre-miRNAs are; mir156a, 156b, 156e, 156f, 156i, 156k, 157b, 159, 159a, 159b, 159e, 160, 160a, 160b, 160c, 160d, 162, 162a, 162b, 164, 164a, 164d, 164b, 165a, 167a, 166a, 167b, 166b, 167c, 166d, 167d, 166c, 167h, 169, 168, 169a, 168a, 169b, 168b, 169e, 169g, 169j, 169l, 169s, 170, 171e, 171t, 172, 171, 172b, 171a, 172c, 171b, 172d, 171c, 172j, 319, 319a, 319b, 319h, 390, 390a, 390b, 390d, 393a, 393c, 394a, 394b, 395a, 395b, 395e, 395j, 395m, 396, 396a, 396b, 396c, 396k, 398, 398a, 399, 399b, 399d, 399c, 399h, 405b, 408, 408c, 416, 479, 774b, 783, 827, 862, 862a, 864, 867, 1310, 1508a, 1510b, 1513b, 2086, 2108a, 2111b, 2118, 2592av, 2600c, 2607, 2643b, 2936, 3440, 3449, 4223, 4246, 4342, 4363, 4387a, 4390, 4403, 4413b, 4414b, 4992, 4993, 4995, 5019, 5032, 5205d, 5219, 5245, 5276, 5288, 5289b, 5291a, 5293, 5298c, 5299, 5368, 5373, 5643a, 5653, 5654, 5667, 5679, 5773, 5784, 5995b, 6034, and 6299.

According to Bonnet *et al.*, (2004) computational and experimental investigation shows that plant miRNAs and their targets between dicot (*A. thaliana*) and monocot (*O. sativa*) are conserved. Bologna *et al.*, (2009) reported that the pre-miRNAs 159 and 319 are well conserved among dicot and monocot. The conservation of mature and pre-miRNAs in species suggests that they play significant role in origin and evolution of multicellular plant life (Reinhart *et al.*, 2002).

In the same study, majority of the dicot reported pre-miRNAs; 2499 out of 2647 are observed as diverged between dicot and monocot. This is showing pre-miRNAs based evolutionary adjustment by the dicots over monocots.

**Convergence and divergence of monocot with bryophyta and pteridophyta pre-miRNAs:** The 969 pre-miRNAs of monocots were studied for convergence and divergence with the nucleotide database of two plants of bryophyta and pteridophyta (*P. patens* and *S. moellendorffii*). Only 26 out of 969 pre-miRNAs of monocot were found as converged with bryophyta and pteridophyta. These pre-miRNAs are observed with 30% to 100% query coverage. The maximum converged pre-miRNAs are found with 30% to 50% query coverage. More convergence is observed between monocot and *P. patens*. The converged pre-miRNAs are mir156a, 156k, 166c, 166d, 166e, 166f, 168, 169i, 171b, 390, 395c, 396b, 529, 815a, 1848, 2091, 2120, 2275b, 2868, 5157a, 5523, 5789, 5797, 5809, 5819 and 5833.

Kozomara & Griffiths-jones (2011) reported 1505 miRNA families in 58 species of monocots with bryophyta and pteridophyta. The five miRNAs (miR156, 160, 165, 166, 167 and 319) are found in monocot, dicot, lycopode, mosses, and gymnosperm, thus it conserved in all embryophyta (Yang *et al.*, 2007). In this project monocot pre-miRNAs, the 943 of 969 are observed as diverged with bryophyta and pteridophyta.

**Convergence and divergence between monocot and**

**gymnosperm pre-miRNAs:** For this study, 969 pre-miRNAs were subjected for homology search against the nucleotide database of three species of gymnosperm. The only 27 out of 969 pre-miRNAs were found as converged between monocot with gymnosperm. These pre-miRNAs are observed between 30% to 100% query coverage. The maximum converged pre-miRNAs between monocot and gymnosperm were found with 30% to 50% query coverage. More convergence is observed between monocot and *P. taeda*. The converged pre-miRNAs are mir156b, 159a, 166c, 166d, 166h, 171b, 171k, 172a, 393b, 395d, 395q, 396c, 397a, 399g, 408b, 414, 834, 1428e, 1846e, 2093, 2863b, 2868, 5072, 5074, 5523, 5819 and 5833.

When some small RNAs sequences of *O. sativa* and *P. taeda* were aligned with their parent EST sequences in some distinct regions, a shared of characteristics in these libraries with known miRNA sequences were found. The 336 EST sequences had alignment patterns of small RNAs that judged same to pre-miRNAs and put ahead for folding examination. This method identified new miRNA genes as well as three clusters related to miRNA 166, 396 and 159 families. So they are conserved in both species (Ryan *et al.*, 2008). In the current study, most divergence is observed in the maximum reported monocot pre-miRNAs with gymnosperm that is 942 of 969.

**Convergence and divergence between monocot and dicot pre-miRNAs:** The 969 pre-miRNAs of monocot were studied for convergence and divergence against the database of twenty seven species of dicot. Out of 969 only 157 pre-miRNAs were found converged from monocot with dicot. These pre-miRNAs are observed with 30% to 100% query coverage. The maximum converged pre-miRNAs between monocot and dicot were found with 30% to 50% query coverage. More convergence is observed between monocot and *G. hirsutum*. The converged pre-miRNAs are mir159, 156a, 159a, 156b, 159b, 156d, 159d, 156j, 160a, 162a, 160b, 160e, 160c, 162, 160d, 162b, 164c, 166d, 166e, 167a, 167c, 167d, 167h, 168, 168a, 168b, 169a, 169b, 169e, 169g, 169h, 169m, 169p, 171a, 171b, 171d, 171h, 171i, 171k, 172a, 172c, 172d, 319a, 319b, 390, 393a, 394, 394a, 394b, 395b, 395f, 395g, 395h, 395m, 395n, 395o, 396a, 396b, 396c, 397, 397a, 398a, 398b, 399c, 399e, 399g, 399i, 408, 408b, 414, 417, 419, 426, 437, 438, 444b, 444d, 444f, 528, 528b, 529a, 815a, 816, 818b, 827b, 1318, 1320, 1425, 1432, 1436, 1846b, 1846c, 1848, 1854, 1866, 1871, 1878, 1882b, 2090, 2091, 2093, 2094, 2095, 2098, 2099, 2100, 2102, 2103, 2105, 2118g, 2275d, 2864, 2868, 2870, 2871a, 2924, 2926, 3981, 5071, 5083, 5146, 5157a, 5338, 5381, 5389, 5385, 5492, 5503, 5506, 5522, 5530, 5538, 5539, 5543, 5564b, 5570, 5789, 5797, 5802, 5806, 5809, 5818, 5819, 5820, 5523, 5825, 5828, 6221, 6222, 6223, 6232a, 6233, 6234a, 6252, 6253 and 6255.

The findings of Yong *et al.*, (2013) are also similar to our results, they experimentally proved that 24 miRNA families stem loop structures and mature sequences are conserved in *phoenix dactylefera* (monocot) and seven other plants like, *A. thaliana*, *G. max*, *Malus domestica*, *Solanum lycopersicum*, *Malus domestica*, *O. sativa* and *populus tricoarpa*. The miRNA families are mir159,

156, 157, 168, 171, 319, 170, 160, 167, 166, 164, 172, 169, 398, 399, 390, 162, 395, 397, 394, 393, 482, 396, and 408. In the miRBase (Release21: June, 2014) these 24 miRNA sequences are available from 19 species.

While more divergence between the pre-miRNAs of dicot and monocot are observed. From the maximum reported monocot pre-miRNAs that are 812 out of 969 are observed as diverged between monocot and dicot.

**Convergence and divergence of gymnosperm with bryophyta and pteridophyta pre-miRNAs:** 104 pre-miRNAs of gymnosperm were studied for convergence and divergence with the nucleotide database of two plants of the bryophyta and pteridophyta (*P. patens* and *S. moellendorffii*). Only eight out of 104 pre-miRNAs were found as converged from gymnosperm with bryophyta and pteridophyta. These pre-miRNAs were observed with 30% to 90% query coverage. The maximum converged pre-miRNAs between gymnosperm with bryophyta and pteridophyta were found with 30% to 50% query coverage. More convergence is observed between gymnosperm and *S. moellendorffii*. The converged pre-miRNAs are mir156b, 166a, 166b, 396, 396a, 396b, 950a and 1310.

Yang *et al.*, (2007) also reported nine miRNA families (mir156, 159, 160, 164, 166, 165, 167, 169, 172, and 390) among gymnosperm, pteridophytes, bryophytes and angiosperms. These nine miRNAs in angiosperm involved in flower development. This conservation was checked for 58 plants; in which one species was selected from bryophytes, two from pteridophytes, four from gymnosperm, 41 from dicot and ten plants from monocot.

Here, most of the reported gymnosperm pre-miRNAs; 96 out of 104 were observed as diverged with the bryophyta and pteridophyta. This means difference in pre-miRNA sequences is required to make difference between gymnosperm, bryophyta and pteridophyta.

**Convergence and divergence between gymnosperm and monocot pre-miRNAs:** Total 104 pre-miRNAs were studied for convergence and divergence with three plants of the monocot. Only 20 out of 104 were found as converged pre-miRNAs from gymnosperm with monocot. These pre-miRNAs were observed with 30% to 100% query converge. The maximum converged pre-miRNAs between gymnosperm with monocot were found with 30% to 50% query converge. More convergence is observed between gymnosperm and *O. sativa* of monocot. The converged pre-miRNAs are mir156b, 159a, 160a, 160b, 166a, 166b, 166c, 390, 396, 396a, 396c, 408, 482b, 482d, 951, 1310, 3706, 3708, 3709a and 3710.

The microarray based study was carried on 2521 miRNAs of 43 plant species from monocot (*O. sativa* and *Z. mays*), gymnosperm (*P. abies* and *P. taeda*), algae and dicot. About 99 miRNAs from 22 miRNA families miR156, 159, 164, 165, 166, 167, 168, 171, 319, 396, 397, 399, 528, 529, 535, 894, 1432, 1450, 1561, 2919, 2936, and 3946, are identified through microarray in the leaf tissue of *Pelliaepeda tisecta*, and the 1957 sequences were observed in complementary with 43 species of plants (Bo *et al.*, 2012).

Whereas, the maximum reported gymnosperm pre-miRNAs; 84 of 104 pre-miRNAs were observed as diverged between gymnosperm and monocot. Showing an

indication of pre-miRNAs based evolution from gymnosperm to monocot.

**Convergence and divergence between gymnosperm and dicot pre-miRNAs:** To investigate Convergence and divergence of pre-miRNAs between gymnosperm and dicot, 104 pre-miRNAs of gymnosperm were studied. Only 26 out of 104 pre-miRNAs were found as converged from gymnosperm with dicot. These pre-miRNAs were observed with 30% to 100% query converge. The maximum converged pre-miRNAs between gymnosperm with dicot were found with 30% to 50% query converge. More convergence was observed between gymnosperm and *G. hirsutum* of dicot. The converged pre-miRNAs are mir156a, 156b, 159c, 160a, 166a, 166b, 166c, 169, 171, 390, 395, 396, 396a, 396b, 396c, 408, 482a, 482d, 1310, 1314, 3694, 3697, 3699, 3700, 3710 and 3712.

There are two types of miRNAs, the young and ancient miRNA. The ancient miRNAs are conserved evolutionarily and express highly. Whereas the young miRNAs are expressed less or may express in particular conditions and are found in few species (Taylor *et al.*, 2014). There are several ancient miRNAs that are found universally in all land plants. These miRNA are miR156, 160, 165, 166, 167, 319, 390, 395, and 408. The diverged miRNAs are miR822 and 839 in *A. thaliana*, miR7695 in *O. sativa* (Taylor *et al.*, 2014). In this research majority of the reported gymnosperm pre-miRNAs that are 78 out of 104 were observed as diverged between gymnosperm and dicot.

**Convergence and divergence of bryophyta and pteridophyta with gymnosperm pre-miRNAs:** The 287 pre-miRNAs of bryophyta and pteridophyta were studied for convergence and divergence with the three plants of the gymnosperm. Only five out of 287 pre-miRNAs of bryophyta and pteridophyta were found as converged with gymnosperm. These pre-miRNAs are observed with 30% to 50% and 70% to 90% query coverage. The maximum converged pre-miRNAs between bryophyta and pteridophyta with gymnosperm were found with 30% to 50% query converge convergence. More convergence is observed of bryophyta and pteridophyta with *P. taeda* of gymnosperm. The converged pre-miRNAs are mir156a, 156b, 477c, 1088 and 1102.

The alignment of many land plant pre-miRNAs 159/319 and its phylogenetic tree were also conducted by Bologna *et al.*, (2009). The result of their research showed that the miRNA 159 and 319 had a common origin. In land plants the duplex exterior the miRNA 159 and miRNA 319 are highly conserved in stem loops of miRNA 159/319. Furthermore, the partitions in the conserved area are constant with the phasing of miRNA 159 and 319 from bryophyte to angiosperm. In the current study, the maximum number of bryophyta and pteridophyta pre-miRNAs that are 282 out of 287 are observed as diverged with gymnosperm.

**Convergence and divergence of bryophyta and pteridophyta with monocot pre-miRNAs:** 287 pre-miRNAs of bryophyta and pteridophyta were studied for convergence and divergence with the three plants of the monocot. Only 27 out of 287 were found as converged pre-miRNAs from bryophyta and pteridophyta with

monocot. These pre-miRNAs were observed with 30% to 90% query coverage. The maximum converged pre-miRNAs between bryophyta and pteridophyta with monocot were found with 30% to 50% query coverage. More convergence of bryophyta and pteridophyta pre-miRNAs is observed with *Z. mays* of monocot. The converged pre-miRNAs are mir156a, 156c, 160b, 319d, 390c, 477c, 529f, 533a, 535c, 1030f, 1038, 1044, 1053, 1066, 1082b, 1085, 1086, 1088, 1093, 1098, 1102, 1107, 1115, 1211, 1215, 2084 and 2085.

In *Pellia endiviifolia*; a bryophyte, 311 miRNA families are known, they are also found in other vascular plants. About 25% of the pteridophytes, and bryophytes specific and known miRNA families, 35% gymnosperm specific miRNAs, 15% monocot specific miRNAs, are also reported in bryophyte (*Pellia endiviifolia*) (Kozomara et al., 2014). Whereas the maximum reported bryophyta and pteridophyta pre-miRNAs; 260 out of 287 were observed as diverged with monocot.

**Convergence and divergence of bryophyta and pteridophyta with dicot pre-miRNAs:** To identify the convergence and divergence of pre-miRNAs between bryophyta and pteridophyta with dicot, 287 pre-miRNAs were studied. Only 18 out of 269 pre-miRNAs were found as converged pre-miRNAs from bryophyta and pteridophyta with dicot. These pre-miRNAs were observed with 30% to 70% query coverage. The maximum converged pre-miRNAs between bryophyta and pteridophyta with dicot were found with 30% to 50% query coverage. More convergence for the bryophyta and pteridophyta, is observed with *B. napus* of dicot. The converged pre-miRNAs are mir156a, 156d, 160a, 396, 529g, 902l, 1027a, 1043, 1082b, 1088, 1093, 1098, 1100, 1101, 1102, 1107, 1109 and 1217.

There are also some miRNAs, which are restricted to vascular plants only. Those miRNAs are; miR435, 528 and 812. The miR1097 is found only in *S. moellendorffii*. While the miR538, 894, 904, and 1030 are found only in *P. patens*. According to Cuperus et al., (2011) eight miRNA families were conserved in monocot, dicot, mosses and pteridophytes. One of them is the miRNA 319. Sylwia et al., (2014) confirmed the presence of these eight families of miRNA in mosses and other plants on land. Whereas, the maximum reported bryophyta and pteridophyta pre-miRNAs; 269 out of 287 were observed as diverged with dicot.

## Conclusion

A total of 4007 known pre-miRNAs from 35 plant species belonging to dicot (2647 from 27 spp), monocot (969 from 3 spp), gymnosperm (104 from 3 spp), bryophyta and pteridophyta (287 from 2 spp) were subjected for convergence and divergence studies. Subsequently, 52% (of 104), 22% (of 969), 11% (of 287) and 10% (of 2647) were found converged among gymnosperm, monocot, bryophyta and pteridophyta, and dicot respectively. The findings from this study would serve as baseline data to make simplicity from the

complexities of convergence and divergence among pre-miRNAs in plant kingdom.

## References

- Altschul, S., W. Gish, W. Miller, E. Myers and D. Lipman. 1990. Basic local alignment search tool. *J. of Mol. Bio.*, 215: 403-410.
- Arazi, T., M. Talmor-Neiman, R. Stav, M. Riese, P. Huijser, and D.C. Baulcombe. 2005. Cloning and characterization of micro-RNAs from moss. *The Plant J.*, 43(6): 837-848.
- Babiarz, J.E., J.G. Rub, Y. Wang, D.P. Bartel and R. Blelloch. 2008. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev.*, 22: 2773-785.
- Baloch, I. A., M. Y. K. Barozai and M. Din. 2013. MicroRNAs: the mega regulators in eukaryotic genomes. *Pure and Appl. Biol.*, 2(3): 83-88.
- Baloch, I.A., M.Y.K. Barozai and M. Din. 2015a. Identification and characterization of 25 microRNAs and their targeted proteins in apricot (*Prunus armeniaca* L.). *J. Ani. & Plant Sci.*, 25(5): 1466-1476.
- Baloch, I.A., M.Y.K. Barozai, M. Din and A.K.K. Achakzai. 2015b. Computational identification of 18 microRNAs and their targets in three species of rose. *Pak. J. Bot.*, 47(4): 1281-1285.
- Barozai, M.Y.K. 2012a. Identification and characterization of the microRNAs and their targets in *Salmo salar*. *Gene.*, 499(1): 163-168.
- Barozai, M.Y.K. 2012b. The novel 172 sheep (*Ovis aries*) microRNAs and their targets. *Mol. Biol. Rep.*, 39(5): 6259-6266.
- Barozai, M.Y.K. 2012c. The MicroRNAs and their targets in the channel catfish (*Ictalurus punctatus*). *Mol. Biol. Rep.*, 39(9): 8867-8872.
- Barozai, M.Y.K. 2012d. Insilico identification of microRNAs and their targets in fiber and oil producing plant Flax (*Linum usitatissimum* L.) *Pak. J. Bot.*, 44(4): 1357-1362.
- Barozai, M.Y.K. 2013. Identification of microRNAs and their targets in *Artemisia annua* L. *Pak. J. Bot.*, 45(2): 461-465.
- Barozai, M.Y.K. and A.H. Wahid. 2012. Insilico identification and characterization of cumulative abiotic stress responding genes in Potato (*Solanum tuberosum* L.). *Pak. J. Bot.*, 44(SI): 57-69.
- Barozai, M.Y.K. and M. Din. 2017. Initial screening of plant most conserved MicroRNAs targeting infectious viruses: HBV and HCV. In Applied Sciences and Technology (IBCAST), 2017 14<sup>th</sup> International Bhurban Conference on (pp. 192-196). IEEE.
- Barozai, M.Y.K. and M. Din. 2014. The relationship between codon usage bias and cold resistant genes. *Pak. J. Bot.*, 46(3): 823-826.
- Barozai, M.Y.K. and T. Husnain. 2011. Identification of biotic and abiotic stress up-regulated ESTs in *Gossypium arboreum*. *Mol. Bio. Rep.*, 39(2): 1011-1018.
- Barozai, M.Y.K., A.G. Kakar and M. Din. 2012a. The relationship between codon usage bias and salt resistant genes in *Arabidopsis thaliana* and *Oryza sativa*. *Pure App. Biol.*, 1(2): 48-51.
- Barozai, M.Y.K., F. Bashir, S. Muzaffar, S. Afzal, F. Behlil and M. Khan. 2014a. In-silico identification and characterization of organic and inorganic chemical stress responding genes in yeast (*Saccharomyces cerevisiae*). *Gene.*, 550(1): 74-80.
- Barozai, M.Y.K., I.A. Baloch and M. Din. 2011. Computational identification of MicroRNAs and their targets in two species of evergreen Spruce tree (*Picea*). *P. WASET.*, 75: 413-418.

- Barozai, M.Y.K., I.A. Baloch and M. Din. 2012b. Identification of MicroRNAs and their targets in *Helianthus*. *Mol. Bio. Rep.*, 39(3): 2523-2532.
- Barozai, M.Y.K., M. Irfan, R. Yousaf, I. Ali, U. Qaisar, A. Maqbool, M. Zahoor, B. Rashid, T. Hussnain and S. Riazuddin. 2008. Identification of micro-RNAs in cotton. *Plant Phy. & Biochem.*, 46(8-9): 739-51.
- Barozai, M.Y.K., S.Q. Shah, M. Din and R. Muhammad. 2014b. Codon usage bias and RNA secondary structures analysis for virus resistant genes in *Arabidopsis thaliana* and *Oryza sativa*. *Pure Appl. Biol.*, 3(2): 81-91.
- Bartel, D.P. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 116: 281-297.
- Behlil, F., M. Y. K. Barozai, F. Bashir, S. Muzaffar, S. Afzal and M. Khan. 2016. Profiling and characterization of chemical and pathogenic cumulative stress responding genes in mouse liver. *Pure and Appl. Biol.*, 5(4): 1323-1333.
- Bibi, F., M.Y.K. Barozai and M. Din. 2017. Bioinformatics profiling and characterization of potential microRNAs and their targets in the genus *Coffea*. *Turkish Journal of Agri. and Forest.*, 41(3): 191-200.
- Bo, W., D. Miao, C. Wenduo, L. Xuefeng, F. Ruijuan and X. Tao. 2012. Microarray identification of conserved microRNAs in *Pinellia pedatisecta*. *J. Gene.* 498: 36-40.
- Bologna, N.G., J.L. Mateos, E. G. Bresso and J.F. Palatnik. 2009. A loop-to-base processing mechanism underlies the biogenesis of plant microRNAs miR319 and miR159. *The EMBO J.*, 28(23): 3646-3656.
- Bonnet, E., J. Wuyts, P. Rouze and Y. Van-de-Peer. 2004. Detection of 91 potential conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* identifies important target genes. *PNAS USA*, 101: 11511-11516.
- Cuperus, J.T., N. Fahlgren and J.C. Carrington. 2011. Evolution and functional diversification of MIRNA genes. *The Plant Cell*, 23(2): 431-442.
- Din, M. and M.Y.K. Barozai. 2014a. Profiling and characterization of eggplant (*Solanum melongena* L.) microRNAs and their targets. *Mol. Biol. Rep.*, 41(2): 889-894.
- Din, M. and M.Y.K. Barozai. 2014b. Profiling microRNAs and their targets in an important fleshy fruit: Tomato (*Solanum lycopersicum*). *Gene.*, 535(2): 198-203.
- Din, M., M.Y.K. Barozai and I.A. Baloch 2016. Profiling and annotation of microRNAs and their putative target genes in chilli (*Capsicum annuum* L.) using ESTs. *Gene Reports*, 5: 62-69.
- Din, M., M.Y.K. Barozai and I.A. Baloch. 2014. Identification and functional analysis of new conserved microRNAs and their targets in potato (*Solanum tuberosum* L.). *Turk. J. Bot.*, 38(6): 1199-1213.
- Griffiths-Jones, S. 2004. The microRNA Registry. The microRNA Registry. *Nucleic Acids Res.*, 32D: 109-111.
- Gul, Z., M.Y.K. Barozai and M. Din 2017. In-silico based identification and functional analyses of miRNAs and their targets in Cowpea (*Vigna unguiculata* L.). *Aims Genet.*, 4(2): 138-165.
- Jahan, S., M. Y. K. Barozai, M. Din, H. Achakzai and A. Sajjad. 2017. Expressional studies of microRNAs in Hepatitis B Patients of Quetta, Pakistan. *Pure and Appl. Biol.*, 6(3): 1044-1052.2004.
- Jones-Rhoades M.W., D.P. Bartel and B. Bartel. 2006. MicroRNAs and their regulatory roles in plants. *Annu. Rev. Plant Biol.*, 57: 19-53.
- Kozomara, A. and S. Griffiths-Jones. 2011. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucl. Acids Res.*, 39.
- Kozomara, A., S. Hunt, M. Ninova, S. Griffiths-Jones and M. Ronshaugen. 2014. Target repression induced by endogenous microRNAs: large differences, small effects. *PLoS One*, 9(8): 104286.
- Lee, R.C., R.L. Feinbaum and V. Ambros. 1993. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, 75: 843-54.
- Luo Y, Z. Guo and L. Li. 2013. Evolutionary conservation of microRNA regulatory programs in plant flower development. *Dev. Biol.*, 380: 133-144.
- Reinhart, B.J., E.G. Weinstein, M.W. Rhoades, B. Bartel and D.P. Bartel. 2002. MicroRNAs in plants. *Genes Dev.*, 16: 1616-1626.
- Ryan, D., A. Morin, D. Gozde, H.E. Elena, M. Alexander, R. Vincent, S. Elaine, S. Mardis, Cenk and J.U. Peter. 2008. Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa*. *Genome Res.*, 18(4): 571-584.
- Shah, S.Q., M. Din, M.Y.K. Barozai and A.K.K. Achakzai. 2017. Codon usage bias analysis for abiotic stress genes in *Arabidopsis thaliana* and *Oryza sativa*. *Pak. J. Bot.*, 49: 15-23.
- Sylwia, A., P. Pawel, P. Halina, M.P. Andrzej, S. Izabela and Przemyslaw. 2014. The liverwort *Pellia endiviifolia* shares microtranscriptomic traits that are common to green algae and land plants. *New Phytologist*, 206: 352-367.
- Taylor, R.S., J. E. Tarver, S.J. Hiscock and P.C. Donoghue. 2014. Evolutionary history of plant microRNAs. *Trends in plant sci.*, 19(3): 175-182.
- Voinnet, O. 2009. Origin, biogenesis, and activity of plant microRNAs. *Cell*, 136: 669-687.
- Yang, T., L. Xue and L. An. 2007. Functional diversity of miRNA in plants. *Plant Sci.*, 172(3): 423-432.
- Yong, X., X. Wei, Y. Yaodong, S.M. Annaliese, L. Xintao and M. Zilong. 2013. Characterization and Evolution of Conserved MicroRNA through Duplication Events in Date Palm (*Phoenix dactylifera*). *PLoS One*, 8(8): 71435.
- Zhang, B.H., X.P. Pan, G.P. Cobb and T.A. Anderson. 2006. Plant microRNA: a small regulatory molecule with big impact. *Dev. Biol.*, 289: 3-16.

(Received for publication 18 January 2017)