Apple *CALCINEURIN B-LIKE PROTEIN10* genes have evolved to be novel targets of miR167s through sequence variation

Ashutosh Kumar and Ananda K. Sarkar*

National Institute of Plant Genome Research, Aruna Asaf Ali Marg, New Delhi 110 067, India

The miR167s and its target *ARF6/8* are relatively conserved among diverse plant species and have been implicated in reproductive and root development in *Arabidopsis*. Here we show that some of the *CNBL* family members have evolved to be targets of miR167s in apple. Despite strong conservation between apple and *Arabidopsis* CNBLs, *AtCNBLs* are not miR167 targets. The sequence variation in apple-miR167a and *MdCNBLs* has created target sites for apple-miR167a in *MdCNBL10*s. Therefore, we suggest that during the course of evolution, natural selection through sequence variation played a crucial role by choosing different targets among plant species for the same miRNA.

Keywords: *Arabidopsis thaliana*, *CNBL10*, miR167, *Malus domestica* (apple), miRNA evolution.

THE diversity among organisms having the same genetic material arises due to the variation in functionality of their molecular mechanism. The molecular mechanism of biological processes involves dynamic network of genes responsible for multiple functions. Besides protein coding genes, small non-protein coding RNAs have recently been shown to be crucial components of gene regulatory network. These small RNAs (sRNAs, 19–24 bp long) regulate diverse developmental processes, including root and shoot growth[1], by transcriptional or post-transcriptional gene silencing mechanisms[2]. The two sRNA classes, microRNAs (miRNAs) and small interfering RNAs (siRNAs) majorly comprise sRNA population in plants[3]. The miRNAs act either by cleaving target mRNA(s) by pairing with it at complementary sequence to direct post-transcriptional gene silencing (PTGS) or by inhibiting translation of mRNAs[3,4]. It is generally considered that miRNAs and their target(s) with perfect or near-perfect complementary sites are conserved among various plant species[5]. Similarly, we have recently shown that miR167s, which are required for flower development through negative regulation of *AUXIN RESPONSE FACTOR* (*ARF6/8*) transcript, have undergone thorough co-evolution with their target *ARF6/8*s among diverse plant species[6,7]. We found a novel non-*ARF* target of mdm-miR167, *CALCINEURIN B-LIKE10* (*CNBL10*) in apple (*Malus domestica*), which was experimentally validated through 5′ RLM-RACE PCR[6]. Moreover, our study has identified and predicted non-*ARF6/8* target in crops such as *Oryza sativa*, *Glycine max*, etc[6]. This suggests that sequence variation in both miRNAs and complementary sequence of targets occasionally leads to functional diversification of miRNAs, which is not well addressed. Conventionally, it is assumed that like miRNAs, targets are also conserved among diverse plants species because functionality of miRNAs is depicted through their targets. Changes that occur during the course of evolution in both the mature miRNA sequence and their complementary regions of target genes may lead to alteration in that miRNA-mediated gene regulation and functional diversification. Therefore, for better understanding the regulatory role of a biologically important miRNA, it is prerequisite to understand whether their functional diversification happened through sequence variation in miRNA and/or target sites. In this work, we extended our recent study where we speculated that miRNAs and their targets co-evolved, which played an important role for their functional divergence, and hypothesized the possible cause of the functional divergence of mdm-miR167 in *Malus domestica*.

The miR167 sequences of *Arabidopsis* and apple were retrieved from miRNA registry database (miRBase version 21, http://microrna.sanger.ac.uk/). It was found from our recent study, while performing multiple sequence alignment, that the ath-miR167a, b and mdm-miR167b-g have identical sequences, which therefore clustered and were named together as unique miR167-1 (UmiR167-1), whereas ath-miR167c and d have difference in their sequences. Similarly, mdm-miR167a has difference in its sequence and did not cluster with other mdm-miR167s and ath-miR167s, therefore was named as unique miR167-2 (UmiR167-2)[6]. The target for the UmiR167-1 was predicted with the help of *psRNATarget* tool[8], as *ARF6*, which was already reported[7]. The *psRNATarget* tool was used to identify or predict a novel target gene for UmiR167-2, which has been validated through 5′ RLM-RACE PCR[6]. The nucleotide and protein sequences of *CNBL* family members of *Arabidopsis* and apple were retrieved from TAIR (version 10) and NCBI respectively (Table 1). The nomenclature for the species specific *CNBL* genes was done by putting the prefix 'At' for *Arabidopsis* and 'Md' for apple. The gene names *CNBL* and *CBL* are synonymous; the entries of protein sequences in UniProt database (http://www.uniprot.org/uniprot) are indicated as 'CNBL', whereas TAIR indicated these genes as '*CBL*'. These sequences were used to check if they were possible target of UmiR167-1 or UmiR167-2 by using *psRNATarget* tool. Homology between *AtCNBLs* and *MdCNBLs* was identified using phylogenetic analysis tool MEGA6 keeping default parameter settings[9]. Further, to reveal critical sequence

*For correspondence. (e-mail: aksarkar@nipgr.ac.in)

Table 1. ***a***, Nucleotide and protein sequences of *CNBL* family members in *Arabidopsis*

Gene*	Notation*	Protein accession number	Nucleotide accession number	TAIR accession number
AtCBL1	*AtCNBL1*	AAC26008	AF076251	At4g17615
AtCBL2	*AtCNBL2*	AAC26009	AF076252	At5g55990
AtCBL3	*AtCNBL3*	AAC26010	AF076253	At4g26570
AtCBL4	*AtCNBL4*	AAG28402	AF192886	At5g24270
AtCBL5	*AtCNBL5*	AAG28401	AF192885	At4g01420
AtCBL6	*AtCNBL6*	AAG28400	AF192884	At4g16350
AtCBL7	*AtCNBL7*	AAG10059	AF290434	At4g26560
AtCBL8	*AtCNBL8*	AAL10300	AF411957	At1g64480
AtCBL9	*AtCNBL9*	AAL10301	AF411958	At5g47100
AtCBL10	*AtCNBL10*	AAO72364	AF490607	At4g33000

b, Nucleotide and protein sequences of CNBL family members in apple

Gene*	Notation*	Protein accession number	Gene accession number
MdCBL1	*MdCNBL1.1*	XP_008373643	LOC103436960
	MdCNBL1.2	XP_008363760	LOC103427483
	MdCNBL1.3	XP_008356465	LOC103420179
MdCBL2	*MdCNBL2*	XP_008358898	LOC103422620
MdCBL3	*MdCNBL3.1*	XP_008384609	LOC103447204
	MdCNBL3.2	XP_008355026	LOC103418695
	MdCNBL3.3	XP_008344727	LOC103407603
MdCBL4	*MdCNBL4.1*	XP_008382769	LOC103445526
	MdCNBL4.2	XP_008376721	LOC103439871
	MdCNBL4.3	XP_008354255	LOC103417881
MdCBL5	N/A	–	–
MdCBL6	N/A	–	–
MdCBL7	*MdCNBL7.1*	XP_008343454	LOC103406228
	MdCNBL7.2	XP_008341705	LOC103404551
	MdCNBL7.3	XP_008337551	LOC103400660
MdCBL8	N/A	–	–
MdCBL9	N/A	–	–
MdCBL10	*MdCNBL10.1*	XP_008391317	LOC103453555
	MdCNBL10.2	XP_008377510	LOC103440590
	MdCNBL10.3	XP_008365025	LOC103428675
	MdCNBL10.4	XP_008357234	LOC103420977

*CNBL and *CBL* are synonymous; the entries of protein sequences in UniProt database are 'CNBL', whereas genes indicated as '*CBL*' in TAIR.

variation between the complementary UmiR167 sequences and their target genes was analysed through ClustalX(2.1)[10].

The novelty in the function of plant miRNAs crucially depends on how they evolved, which includes either changes in their functional mature sequences, processing and expression pattern. In apple, the *MdCNBL10* was shown to be the novel target of mdm-miR167a, which is exceptionally processed differently (from 3′ end of the precursors), and had undergone sequence diversification[6]. Therefore, it was important to find out the possible diversification in the functionality of these miR167s and *CNBL* genes. To study whether other members of *CNBL* gene family were also targeted by the UmiR167-2, we retrieved all the sequences from *Arabidopsis* and apple. In *Arabidopsis*, the *AtCNBL* gene family consisted of 10 members (*AtCNBL1-10*), whereas in apple, 6 members were found (except *MdCNBL5*, *6*, *8* and *9*, which were missing), while some of them had multiple copies (Table 1). Our study identified the homology between AtCNBL and MdCNBL, using the MEGA6 tool (Figure 1 *a*). We observed that the available members of MdCNBL were orthologous to the same members of AtCNBL, except MdCNBL7.1, 7.2 and 7.3. MdCNBL7.1 and 7.3 clustered with AtCNBL8, whereas MdCNBL7.2 clustered with AtCNBL5 (Figure 1 *a*). All these members of *CNBL* gene family were used for checking complementarity with UmiR167-2, using *psRNATarget* tool, by keeping default parameter settings. Interestingly, the *MdCNBL10* has four subfamily members – *MdCNBL10.1*, *MdCNBL10.2*,

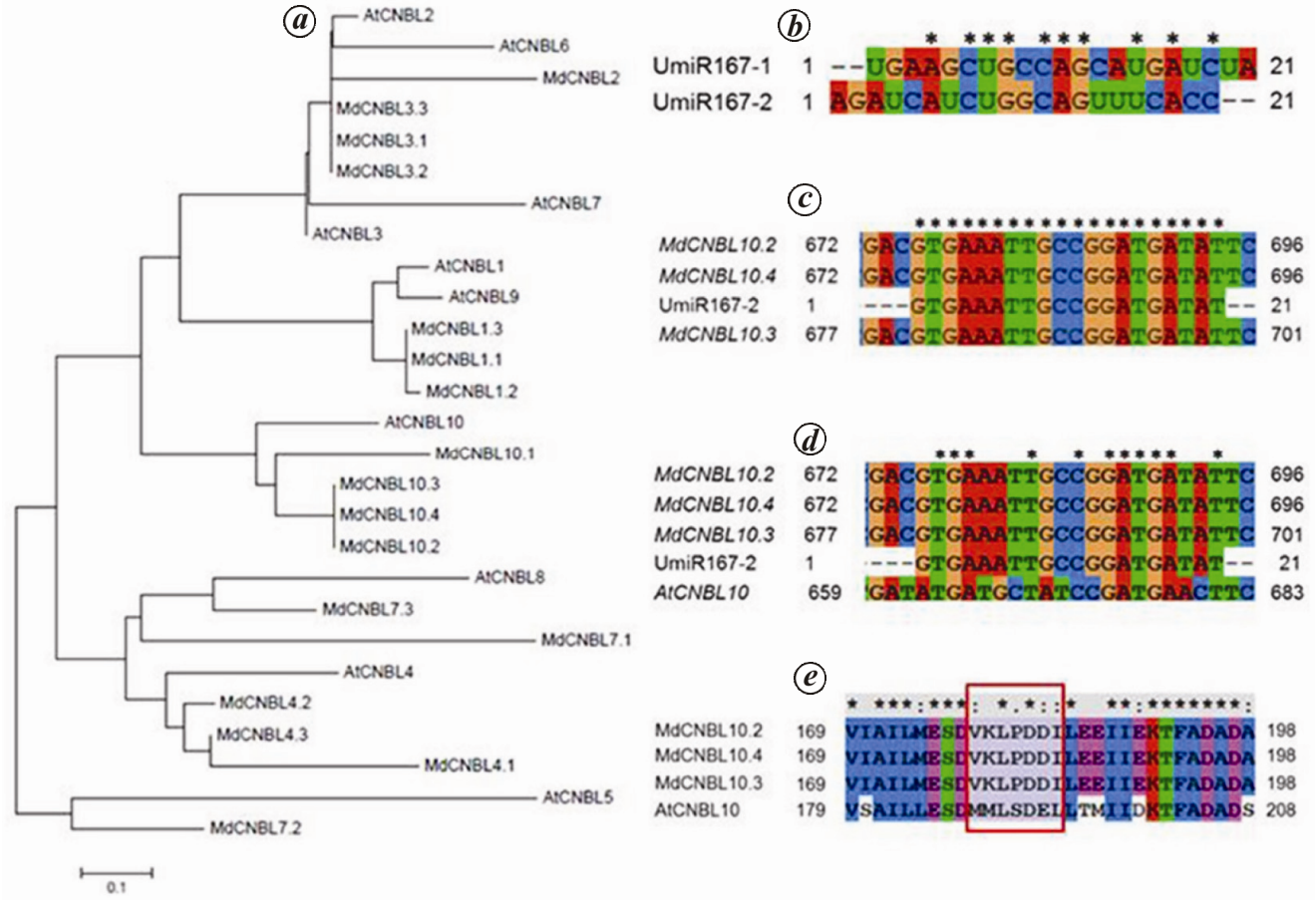

Figure 1. ***a***, Homology identification between AtCNBL and MdCNBL; ***b***, Pairwise alignment between UmiR167-1 and UmiR167-2; ***c***, MSA between the targets *MdCNBL10.2*, *10.3* and *10.4* and complementary sequence of UmiR167-2; ***d***, MSA among *AtCNBL10*, *MdCNBL10.2, 10.3*, *10.4* and complementary sequence of UmiR167-2; ***e***, MSA among MdCNBL10.2, MdCNBL10.3, MdCNBL10.4 and AtCNBL10; red-coloured rectangular box represents amino acids encoded by complementary sequence of UmiR167-2.

MdCNBL10.3 and *MdCNBL10.4*, of which *MdCNBL10.1* was closest to *AtCNBL10.1*; the remaining three members clustered together (Figure 1 *a*; Table 1). We observed that only *MdCNBL10.2*, *MdCNBL10.3* and *MdCNBL10.4* mRNAs were having binding site for UmiR167-2, indicating them as targets of mdm-miR167a. We have recently proved though 5′ RLM-RACE PCR that *MdCNBL10.3* mRNA was cleaved by mdm-miR167a (ref. 6). Our detailed sequence analysis and phylogenetic study of *MdCNBL10.2*, *MdCNBL10.3* and *MdCNBL10.4* shows perfect sequence identity including miR167 binding sites. Therefore, this suggests that our previously validated target *MdCNBL10* included all these three targets. Though *AtCNBL10* was orthologue of *MdCNBL10*, it did not show complementarity with UmiR167-2. Belonging to the same class of miRNA, UmiR167-1 and UmiR167-2 were having different range of targets. The UmiR167-1 was showing complementarity only with *AtARF6* and *AtARF8*, while UmiR167-2 was only complementary to *MdCNBL10.2*, *MdCNBL10.3* and *MdCNBL10.4*. The sequence alignment of both UmiRNAs with the help of ClustalX(2.1) showed a colossal difference between the sequences, only showing ~40% identity (Figure 1 *b*). This suggests that the ability of UmiR167-1, UmiR167-2 to target *ARF6/8* and non-*ARFs* respectively, is caused by their sequence variation in the course of evolution. Our recent study showed that miRNA such as miR166s, despite their sequence conservation, shows sequence variation in target complementary sites leading to the functional diversification in moss (*Physcomitrella patens*), where ppt-miR166m targets non-*HD-ZIPIII* transcript[11]. As the sequence of UmiR167-2 was significantly different from UmiR167-1, it was not showing sufficient complementary binding with *ARF6/8*. The level of complementarity of a miRNA to its possible targets may also vary due to the sequence variation in the target sites. Therefore, we performed multiple sequence alignment between the targets *MdCNBL10.2*, *10.3* and *10.4* along with complementary sequence of UmiR167-2, which have shown 100% identity (Figure 1 *c*). Further, *AtCNBL10*, an orthologue of *MdCNBL10*, was also included for multiple sequence alignment with UmiR167-2 and *MdCNBL10.2*, *10.3*, *10.4*. Inclusion of *AtCNBL10* during multiple sequence alignment has again lowered

the identity up to 55% (Figure 1 *d*). This suggests that the deviation in complementary miRNA binding site of a target gene depends upon the uniqueness of a miRNA as well as its target site sequences, which lead to the functional divergence. Subsequently, we observed that the UmiR167-2 targets, namely *MdCNBL10.2*, *10.3* and *10.4* encode for identical CNBL10 proteins. However, MdCNBL10.1, the closest homolog of AtCNBL10, codes for related protein but is not targeted by UmiR167-2. When the protein sequences of MdCNBL 10.2, 10.3, 10.4 and AtCNBL10 were aligned, it was found that amino acids at the miRNA binding site were not having 100% identity, having only 2 identical, 3 similar, and 1 less similar amino acid. This indicates that the target MdCNBL10 proteins have diverged from non-target MdCNBL10 proteins through sequence variation in critical miRNA complementary sites (Figure 1 *e*). It implies that during the course of evolution, natural selection played a crucial role by choosing different targets among plant species for the same miRNA. The evolution of *MdCNBL10.2*, *10.3* and *10.4* as a target of mdm-miR167a is an interesting aspect to be understood. The cultivation of apple is done in cool temperate zones with warm days and cool nights, which resembles with cool, desert-like climate; the role of CNBL10 was reported as a calcium sensor, which was involved in the signalling pathway during salt and drought stresses[12]. It is possible that the speciation and climatic adaptation in apple was contributed by the co-evolution of mdm-miR167a and its non-*ARFs CNBL10* targets. Similarly, *IAA-Ala Resistant3* (*IAR3*) was identified as a new target of ath-miR167a, which was cleaved during high osmotic stress[13,14]. This indicates potential role of mdm-miR167a mediated regulation of *MdCNBL10* during stress responses in apple.

Therefore, our findings shed light on the basis of functional diversification of miRNAs, as it is often believed that miRNAs and their targets are conserved. The deviation from the canonical target *ARFs* of a miR167 is due to altered pattern of processing of mdm-miR167a from the 3′ end of the precursor sequences (instead of 5′ end), which further led to the sequence variation in the mature miR167 targeting novel gene *MdCNBL10*.

1. Singh, A., Singh, S., Panigrahi, K. C., Reski, R. and Sarkar, A. K., Balanced activity of microRNA166/165 and its target transcripts from the class III homeodomain leucine-zipper family regulates root growth in *Arabidopsis thaliana*. *Plant Cell Rep*., 2014, **33**(6), 945–953.
2. Arikit, S., Zhai, J. and Meyers, B. C., Biogenesis and function of rice small RNAs from non-coding RNA precursors. *Curr. Opin. Plant Biol*., 2013, **16**, 170–179.
3. Allen, E., Xie, Z., Gustafson, A. M. and Carrington, J. C., microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*, 2005, **121**, 207–221.
4. Llave, C., Xie, Z., Kasschau, K. D. and Carrington, J. C., Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. *Science*, 2002, **297**, 2053–2056.
5. Bonnet, E., Wuyts, J., Rouze, P. and Van de Peer, Y., Detection of 91 potential conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* identifies important target genes. *Proc. Natl. Acad. Sci. USA*, 2004, **101**, 11511–11516.
6. Barik, S. *et al.*, Coevolution pattern and functional conservation or divergence of miR167 and their targets across diverse plant species. *Sci. Rep.*, 2015, **5**, doi:10.1038/srep14611.
7. Wu, M. F., Tian, Q. and Reed, J. W., *Arabidopsis* microRNA167 controls patterns of *ARF6* and *ARF8* expression, and regulates both female and male reproduction. *Development*, 2006, **133**, 4211–4218.
8. Dai, X. and Zhao, P. X., psRNA target: a plant small RNA target analysis server. *Nucl. Acids Res*., 2011, **39**, W155–W159.
9. Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S., MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol*., 2013, **30**, 2725–2729.
10. Larkin, M. A. *et al.*, Clustal W and Clustal X version 2.0. *Bioinformatics*, 2007, **23**, 2947–2948.
11. Barik, S., Sarkar Das, S., Singh, A., Gautam, V., Kumar, P., Majee, M. and Sarkar, A. K., Phylogenetic analysis reveals conservation and diversification of micro RNA166 genes among diverse plant species. *Genomics*, 2014, **103**, 114–121.
12. Kim, B. G. *et al.*, The calcium sensor CBL10 mediates salt tolerance by regulating ion homeostasis in *Arabidopsis*. *Plant J.*, 2007, **52**, 473–484.
13. Kinoshita, N. *et al.*, IAA-Ala Resistant3, an evolutionarily conserved target of miR167, mediates *Arabidopsis* root architecture changes during high osmotic stress. *Plant Cell*, 2012, **24**, 3590–3602.
14. Sarkar, A. K., Karthikeyan, M., Gautam, V., Barik, S. and Das, S. S., Improving the plant root system architecture to combat abiotic stresses incurred by global climate changes. In *Climate Change and Abiotic Stress Tolerance*, Wiley, Germany, 2014, pp. 305–324.

ACKNOWLEDGEMENTS. This work was supported by NIPGR short term fellowship, followed by SERB National Post-doctoral fellowship (PDF/2015/000232) to A.K. and NIPGR internal core grant to A.K.S.

Received 17 March 2016; revised accepted 10 August 2016

doi: 10.18520/cs/v112/i01/147-150