To encourage a disease free healthy life, Mother Nature has gifted mankind medicinal plants. According to the WHO survey 80% of the population living in the developing countries rely almost exclusively on traditional medicines for primary health care (Goyal, 2005). The high cost of western pharmaceuticals put modern health care services out of reach of most of the world's population, thus they rely on traditional medicine and medicinal plants to meet their primary health care needs. Even where modern medical care is available and affordable, many people prefer more traditional practices. This is particularly true for our country India, who have tended to retain ethnic medical practices.

In the last decade, there has been considerable interest in resurrecting medicinal plants in western medicine, and integrating their use into modern medical systems. The reasons for this interest are varied, and include the following characteristics offered by medicinal plants i.e. low cost, drug resistance, limitations of medicine, medicinal value, cultural exchange, increased commercial value.

Angiosperms (flowering plants) were the original source of most plant medicines. Some herbs and spices also come from flowering plants. Stevia is a genus of about 23,600 species of herbs and shrubs in the sunflower family (Asteraceae), native to subtropical and tropical regions from western North America to South America. **S. rebaudiana** Bertoni is one of 154 members of the genus Stevia and one of only two species that produce sweet steviol glycosides (Madan et al., 2010). The plant has a long history of medicinal use by the Guarani (a group of culturally related indigenous people of S. America, having been used extensively by them for more than 1,500 years (Misra et al., 2011). Guarani people of Paraguay used stevia, which they called *ka’ahe’è* ("sweet herb"), as a sweetener in *yerba mate* and other foods, medicinally as a cardiac stimulant and as a treatment for obesity, hypertension, heartburn, and to help lower uric acid levels (Tanvir et al., 2005).

**S. rebaudiana** is a perennial herb which accumulates up to 30% (w/w leaf dry weight) diterpenoid SGs (Cardello et al., 1999). SGs are glucosylated derivatives of diterpenoid alcohol steviol whereas stevioside and rebaudioside A are the major SGs found in stevia. Sweetness indices of these and other related compounds ranges between 30 to 300 times higher than that of sucrose (Crammer and Ikan, 1987) and are used as non-calorific sweeteners in many countries of the world. Moreover, the sweetness is superior in quality to that of sugar in terms of mildness and refreshment (Kinghorn and Soejarto, 1989). In addition to SGs, stevia contains a complex mixture.
of sterols, triterpenoids, essential oils and flavonoids. Leaves of stevia and/or extracts containing primarily stevioside have been approved food additives in China, Argentina, Brazil, Japan, South Korea, Russia and Peru. The US FDA has approved stevia as a safe food additive and also granted GRAS (generally recognised as safe) status to it.

A large flux of research activities have diverted to understand the biosynthesis and possible manipulation of diterpenoids in stevia, particularly SGs which are mainly responsible for its ecological and commercial importance. A control over the regulation of these early and late genes involved in SG biosynthetic pathway can help to manipulate the diterpenoid contents (Kumar et al., 2011, Guleria and Yadav 2013, Guleria et al., 2014). Till date several agrotechniques, bioproduct extraction, phytochemical, biological and toxicological studies have been carried out on S. rebaudiana (Madan et al., 2010).

miRNAs have been extensively studied in recent years, and thousands of miRNA genes in the plant kingdom, from mosses and ferns to higher flowering plants have been computationally predicted and/or experimentally cloned either by traditional genetic approaches or by the recently developed NGS strategy (Meng et al., 2011). miRNAs are 20-24 nucleotide small endogenous nonprotein-coding regulatory RNA sequences that are produced by genes distinct from the genes that they regulate. Evidence provided by Allen et al. and Felippes et al, show that some miRNAs evolved by inverted duplications of target gene sequences, whereas others originated from random sequences that either have self complementarity by chance or sequences that represent highly eroded inverted duplications. Since their discovery, several miRNAs have been computationally and/or experimentally identified and characterized in different species. A number of studies have shown that miRNAs play key roles in regulatory functions of gene expression for most eukaryotes (Sunkar and Jhu 2007; Jagadeeswaran et al., 2009) mainly at the post-transcriptional levels (Chapman et al., 2004; Mallory et al., 2004). Several recent findings have implicated miRNAs in a number of biological mechanisms including leaf (Palatnik et al., 2003), stem (Mallory et al., 2004) and root growth (Subramanian et al., 2009), floral organ identity, control of female gamete formation and reproductive development (Millar and Gublar 2005; Olmedo-Monfil et al., 2010), auxin signaling (Rhoades and Bartel 2004), and biotic and abiotic stress responses (Jagadeeswaran et al., 2009).
Till date only a single study identifying 6 conserved miRNAs have been documented and no research activity to identify novel miRNAs in stevia has been conducted. So, there is a need of extensive study to explore the miRNA repertoire in stevia (Guleria and Yadav, 2011). A sudden inclination of extensive research towards investigation of plant miRNAs is due to their indispensable roles in plant development and adaption. At present more than thousand miRNA genes among diverse plant species have been identified, annotated and some of them have been well characterised (Griffiths-Jones et al., 2008). But still a large number of plant species are unexplored and stevia is one of them. Accurate detection and expression profiling of miRNAs will enable a better understanding of their role in stevia growth and development (Jagadeeswaran et al., 2009, Millar and Gublar, 2005; Rhoades and Bartel, 2004), and could provide insights into miRNA-mediated apomictic gene regulatory mechanisms. The main approaches for miRNA identification have been widely undertaken by computational prediction, direct cloning and sequencing. Until recently, most sequence information including ESTs or Genome Survey Sequences (GSS) used for computational prediction of miRNAs were generated by traditional Sanger sequencing methods (Zhang et al., 2006; Zhou et al., 2008). Compared to highly conserved miRNAs, less- or non-conserved miRNAs are often expressed at lower levels, thus making their detection more daunting using small-scale sequencing. The development of NGS technology has greatly improved the capacity to identify low abundance or tissue-specific miRNAs, and has enhanced the discovery of several conserved, non-conserved or lowly expressed miRNAs through cloning and deep sequencing of sRNA and transcriptome libraries in A. thaliana (Rajgopalan et al., 2006; Fahlgren et al., 2008), Triticum aestivum (wheat; (Yao et al., 2007)), Solanum lycopersicum (tomato; (Moxon et al., 2008)), Oryza sativa (rice; Sunkar and Zhu, 2004), Populus trichocarpa (cotton wood; Sunkar et al., 2005), and Manihot esculenta (Cassava; Amiteye et al., 2011).

The objective of this work was thus to identify and completely catalogue conserved and novel miRNAs in S. rebaudiana. The identification of entire sets of miRNA and, subsequently, their targets will lay the foundation for unravelling the complex miRNA-mediated regulatory networks controlling development and other physiological processes. In our study, we used the stevia EST and nucleotide sequence to predict miRNAs and their targets and identified 100 miRNAs belonging to 34 highly conserved families. Also, we identified 24 novel miRNAs whose precursors were potentially generated from stevia EST and nt sequences. All novel sequences have not been earlier described in other plant species.
Further, the unraveling of sRNA guided circuitry in stevia will enhance the value of its gene and EST information and improve our ability to devise strategies to enhance certain essential features of stevia that are less amenable to functional genomics analysis leading to its enhanced nutritive value. The identification of miRNA and their targets is important not only to help us learn more about the roles of miRNAs in stevia development and physiology but also to provide a framework for further designing RNAi based experiments for regulation of gene expression in this species. The work was initiated with an aim to explore the smRNAome of *S. rebaudiana* for the discovery of conserved and novel miRNAs and prediction of their respective targets.

### 5.1 Protocol for isolation of high quality RNA was developed

*S. rebaudiana*, being a rich source of secondary metabolites of medicinal importance, proves to be a good candidate to unravel the miRNA repertoire which regulate the biosynthesis of metabolites such as steviol glycosides and other indispensable processes at gene level. For such molecular studies, the foremost requirement is isolation of good quality of nucleic acids, so emphasis was to obtain RNA of high quality. *S. rebaudiana* contains SGs up to 30% on dry weight basis, complex mixture of sterols, triterpenoids, essential oils and flavonoids. Compounds such as polysaccharides, polyphenols interferes in isolation of good quality RNA. While polyphenols are known to be readily oxidized during nucleic acids extraction to form covalently linked quinones that can irreversibly bind proteins and nucleic acids to form high molecular weight complexes (Loomis, 1974) whereas polysaccharides tend to co-precipitate with nucleic acids in buffers having low ionic strength buffers (Chang et al., 1993). Various protocols have been described for RNA isolation from plant tissues. Previously published protocols and commercial kits were used to isolate sRNA from different parts of plant. Ghawana et al., 2011 and miRNA Easy Spin Kit (Qiagen, Germany) were fused and tried but either quality or quantity or both were very poor especially in case of leaves and pellet was black in color suggesting high amount of polysaccharides co-precipitated with the nucleic acids. Furthermore, gave a very low A (260/280) ratio, A(260/230) ratio and RNA integrity number, so inclusion of RECOIN method was necessary (Singh and Kumar, 2012). RNA isolated by the inclusion of this protocol had improved A (260/280) ratio of 1.88-1.93, A(260/230) ratio of 1.94-2.0, and RNA integrity number of 6.8-8.0 (Fig. 4.2 A, B) with a very clean translucent pellet. In detail, ratio of A260/230 between 1.8 and 2.0 also indicated high purity of RNA with very small polysaccharides and phenolic contamination with inclusion of RECOIN in our RNA isolation. Further the
concentration of RNA samples extracted with inclusion of RECOIN showed significantly higher concentration of 1.647 ng/μl and 1.148 ng/μl whereas RNA sample isolated without inclusion of RECOIN showed low RNA concentration of around 564 ng/μl.

Further, bioanalyzer based analysis suggested high integrity of RNA, RIN = 7.70 and 7.50; (Schroeder et al., 2006) with significant 28S and 18S rRNA peaks isolated by the present protocol whereas very low RIN= 2.30 and nominal 28S and 18S rRNA peaks without inclusion of RECOIN. RNA was then amenable to downstream applications such as sRNA sequencing and library construction.

5.2 Deep sequencing of sRNAs

Due to the important functions of the miRNAs and the recent advances of experimental and computational analytical approaches, interest in these small molecules i.e. miRNAs has increased significantly in recent years. To date, there are 25141 mature miRNA sequences from 193 different species (ranging from viruses to human) collected in the miRBase database (release 19.0). In contrast, the Plant MicroRNA Database (PMRD) contains 10597 miRNAs identified in 127 plants. A great contributor to the amount of recently discovered miRNAs has been the development of high-throughput sequencing methods such as 454 Life Sciences System, Illumina technology and SOLiD system (Motamen, 2010). Traditional strategies like genetic screening, microarray and bioinformatics prediction have miserably failed in species where genome sequence is not available such as stevia. The availability of high throughput NGS technologies has overshadowed this weakness by generating an accurate and comprehensive picture of smRNAome and identification of known miRNAs but also for successful discovery of novel miRNAs with high precision without the need of genome.

To characterize the sRNA transcriptome, the appearance of these three 'deep-sequencing' technologies, have a number of significant advantages in comparison with previous hybridization-based methodologies, such as microarray or PCR-based assays. Firstly, it provides a more integrated view of the miRNAs transcriptome. With the added sequencing depth, high-throughput sequencing have an ability to identify modest or even low abundance miRNAs exhibiting expression differences between different samples, which couldn’t be detected previously. Secondly, direct sequencing also offers the potential to detect variation in mature miRNA length and enzymatic modification of miRNAs. Thirdly, high-throughput
sequencing allows the successful discovery of novel miRNAs, which need not rely on querying candidate regions of the genome but rather can be achieved by direct observation and validation of the folding potential of flanking genomic sequence. Taken together, NGS technologies offer a highly robust, accurate and scalable system that sets a new standard for productively, cost-effectively investigation of sRNA transcriptome (Motamen, 2010).

These technologies have been used in many studies to identify and determine the expression levels of miRNAs that are, for example, conserved, novel, tissue- or developmental stage-specific (Sun et al., 2012, Osanto et al., 2012, Liu et al., 2012, Wang et al., 2012, Joshi et al., 2010). Additionally, experimental miRNA analyses are often supplemented by bioinformatic methods, which are used to process raw sequencing data, predict miRNA genes, precursors, mature sequences and targets, identify isoforms, and classify sRNAs into known miRNA families (Motamen, 2010; Li et al., 2010; Dai et al., 2011). These experimental and computational methods not only allow for low-cost quantitative and qualitative small RNAs analysis but also generate more specific results in shorter time frame (Lhakhang and Chaudhry et al., 2012).

Our research aimed to examine sRNA population in leaves of S. rebaudiana using Illumina platform to identify conserved and novel miRNAs. In this study, we described the first screen for stevia miRNAs by deep sequencing with an aim of gaining insights into various roles of miRNAs. Deep-sequencing can generate millions of short sequences, as in our study using 36 cycled single end sequencing by Illumina genome analyzer II platform (Morozava and Marra, 2008), a total of 30,472,534 sequences were obtained from sRNAome of S. rebaudiana which was parallel with the sRNA sequence analysis findings in other plants. A total of 11,338,273 reads were obtained in early maize seedlings (Wang et al., 2011), 15,702,980 raw sRNA sequence reads were acquired in Chinese fir (Wan et al., 2012) and 22,561,972 sequences in Japanese apricot sRNA library (Gao et al., 2012).

As Illumina platform generates an extensive sRNA library containing 30,472,534 sequences in our case which present substantial informatics challenges for lack of efficient and flexible tools to handle and analysis of such a huge scale of short sequences. All of the common steps in sRNA analysis workflows can be accomplished using freely available software tools. One approach is to use a collection of tools from heterogeneous sources. The advantage of this is that it offers...
complete freedom to choose the most appropriate tool for each step. The steps may be linked together and automated using a scripting language. An alternative approach is to use a comprehensive one-stop-shop. The advantages include convenience of not having to source tools for each step and the reduced reliance on scripting and other computer-based skills that might be limiting for some laboratories (Studholme, 2012).

A good example of such one-stop-shop approach is the UEA sRNA Toolkit which we used in our sRNA library analysis and miRNA prediction. This offered most of the functionality that we required for analysing our sRNA sequencing data. It was accessed via a web server or the underlying software could be downloaded, installed and used locally. But as the standalone tools are currently being substantially revised to improve performance and maintainability UEA web server was preferred (Studholme, 2012). Many of the components of the UEA sRNA Toolkit are essentially high-level wrappers around existing tools for more low-level tasks. For example, it currently uses PatMan (Prufer et al., 2012) to align sRNA sequences against a reference genome and uses the Vienna package (Hofacker et al., 2002) for predicting RNA secondary structures. Some alternatives that provide at least some of the same functionality as the UEA sRNA Toolkit include miRNAkey (Ronen et al., 2010), miRanalyzer (Hackenberg et al., 2009), SeqBuster (Pantano et al., 2010), DSAP (Huang et al., 2010) and mirTools (Zhu et al., 2010), which were not preferred as none of these were primarily intended for plant-derived data.

Further, 17,295,850 high quality sequences were generated from 30,472,534 raw sequences representing 57% of the total raw reads emphasizing the high quality of the generated sRNA library. RNA species 23, 24, 22 and 21 nt in length dominated the sRNA transcriptome with the 23 nt class being the most abundant in our library. This class of 23–25 nt has been named as long miRNAs. This finding is consistent with the working hypothesis that evolutionary change in structure of hairpin can change their affinity for DCL’s and hence the size of miRNAs. So, it is actually DCL3 which is responsible for pri miRNA processing to generate 23 nt long miRNAs in contrast to the highly stringent DCL 1 for 20–21 nt canonical miRNAs. This was consistent with the Watermelon plants grafted on squash (Wm/Sq) at the two true-leaf stage where the most abundant sequences were 23 nt long in length unlike other two grafts grafted onto watermelon (Wm/Wm) or onto bottle gourd (Wm/BG) Wmamong in which, sequences with length of 24 nt were shown to be significant in abundance (Liu et al., 2013). The observation indicated the differential expression pattern of distinct categories of sRNAs in the sample, and differences in the
complexity of sRNA pools indicating different regulation underlying the miRNA-mediated effects of on plant growth and development at different stages of development. Further at times DCL2 can also process certain pri-miRNAs and generate 22 nt long miRNAs (Vazquez et al., 2008) justifying their presence.

In total, 1,752,385 sequences of 24 nt length (11.43%) represented the second highest followed by 22nt sequences (6.83%) and finally 21 nt sequences (5.61%) which is the typical length of plant mature miRNAs. The latter trend of abundances of 24nt sequences followed by 22nt and 21nt was same as observed during seed germination in maize (Wang et al., 2011). Unlike as seen in Brassica oleracea leaves in which the length distribution showed that most of the generated reads had 21 (> 31%), 22 (> 16%), 24 (> 11%) and 23 (> 7%) nucleotides, which are also the most frequent sizes of the known Brassica plant miRNAs (Wang et al., 2012, Korbes et al., 2012, Zhao et al., 2012, Yang et al., 2013).

Additionally, the high throughput sequencing technology provides an alternative way to assess expression profiles of diverse miRNA genes and the number of reads can serve as an index for the relative abundance of diverse miRNAs which was evident from the fact that 2,509,190 sequences out of 15,327,722 sequences (16.75%) were sequenced only once which showed that different miRNAs are present in different quantities and development-stage specific. This ultimately directs that stevia contains a large and diverse sRNA population as was the case seen in C. trifoliate (Song et al., 2010). The sRNA singleton rate of in our study was similar to that of Chinese fir (Wan et al., 2012), A. thaliana (Rajgopalan et al., 2006), O. sativa (Zhu et al., 2008), and P. trichocarpa (Barakat et al., 2007).

5.3 Identification of conserved miRNAs in S. rebaudiana

Previous studies have identified thousands of miRNAs in angiosperms and some of them have been well-characterized (Chen 2004; Lauter et al., 2005). This is the first extensive NGS analysis of miRNAs in S. rebaudiana. A lot of conserved and novel miRNAs were identified. The data might be helpful not only in filling the gap of miRNA registered about basal angiosperm plants but also in understanding the evolution of miRNAs. Further a total of 30, 472, 534 raw sequences generated a profile of 34 conserved miRNA families. Our data were in good harmony with previous studies of miRNA profiling based on exhaustive sequencing of sRNA populations (e.g. in grapevine 24 million reads gave 26 known miRNA families, in peanut 6,009,541 reads gave 22 known miRNA families and in C. trifoliate 13,
106,573 reads gave 42 known miRNA families) (Pantaleo et al., 2010, Zhao et al., 2010, Song et al., 2010). This probably reflects the generally accepted high level of expression reported for conserved miRNAs. Many conserved plant MIR genes arose through extensive genome duplications and rearrangements and thus often have multiple loci. Also, these genes are usually highly expressed, a feature which, until the advent of deep-sequencing technologies, hindered the identification of a second large class of low-to-moderately abundant “younger” miRNAs in plants (Voinnet, 2009).

Stevia miRNA families showed varied abundance from each other. As we employed deep sequencing approach to recover sRNA reads, we used read frequency to measure their abundance. Based on normalized read count per million, conserved miRNAs are divided into 3 groups: abundantly expressed (with read counts greater than 1000) miR159, miR166, miR167, and miR 394; moderately expressed (with read counts 100–1000) miR156, miR168, miR 172, miR319 and miR396 those with low expression (with read counts less than 100) miR164, miR165, miR171, miR393, miR894, and miR 2916. miR160, miR 163, miR 397, miR 403, miR 408, miR 414, miR 482, miR 856, miR 858, miR 1310, miR 1317, miR 1511, miR 1850, miR 2111, miR 3520, miR 5084 and miR 5139 represented by less than 10 normalized reads were the miRNAs with least abundance in stevia. Precisely, the majority of stevia miRNAs were only sequenced less than 1,000 times, and some rare miRNAs were detected less than 10 times, whereas miR167a, miR394a, and miR159a were detected exceptionally high at 9,629, 1554 and 11,759 times respectively. Some were sequenced only a few times, whereas others were present thousands of times, indicating many physiological and biochemical processes are being carried out and stevia contains a large and diverse sRNA population (Wang et al., 2011). Among the 34 miRNA families mir159 showed the largest number of sequenced clones which are in agreement as miR159 is also the most abundant family in Arabidopsis (Rajgopalan et al., 2006). Further mir159 and mir394 with highest abundance in stevia were among the moderately expressed miRNAs in Arachis hypogaea (Zhao et al., 2010) and among the lowly expressed miRNAs in C. trifoliate (Song et al., 2010). In stevia miR156 like wheat (Yao et al., 2007) and rice (Zhu et al., 2008), was among the lowly expressed miRNAs but usually miR156 represents one of the highly abundant miRNA families in diverse plant species e.g. Arachis hypogaea (Zhao et al., 2008), Brachypodium (Zhang et al., 2009), early maize seedlings (Wang et al., 2011) and Medicago truncatula (Qui et al., 2009). The abundance of miR172 was 20 times low as compared to miR156 in our dataset which
is consistent with the previous finding that these two miRNAs are conversely regulated (Wang et al., 2011). Unlike, other plant species where tae-miR169b in wheat and osa-miR169 in rice are the most frequently sequenced miRNAs while in our case miR169b is absent. This may suggest a species-specific expression profile for miRNAs. However, mechanisms causing the differential expression profile of a same miRNA in different plant species is unknown (Wang et al., 2011). Thus this varied abundance of miRNA families suggested that miRNA genes would be differentially transcribed at this young leaves stage.

At one extreme, some miRNA families are conserved in moss, indicating their very ancient origin. These include miR-156, miR-160, miR-319, and miR-390, all of which regulate ancestral transcription factors that specify basic meristem functions, organ polarity and separation, cell division, or hormonal control (reviewed in Garcia, 2008). In addition, the expression of several flowering plant miRNAs was detected in a gymnosperm and a fern, miR160 was detected in a moss, and miR160 was cloned from the moss *Polytrichum juniperinum* demonstrating that many plant miRNAs have remained essentially unchanged since before the emergence of flowering plants (Axtell and Bartel, 2005). Further confirming the fact in our miRNA prediction also, one of the conserved family detected was miR160 family. Other miRNA families evolved after the split between land plants and mosses but before the monocot/dicot divergence. Twenty-one such miRNA families seem universal among angiosperms (reviewed in Axell and Bowman, 2008). Among the 21 miRNA families conserved between dicots and monocots, 5 of which were identified in *S. rebaudiana*. It includes miR159, miR166, miR169, miR171 and miR399. It indicates that the ancient miRNA regulatory system is well-developed in the common ancestors of gymnosperms and angiosperms (Axtell, 2008). Whereas, certain miRNA families are monocot or dicot specific which is further justified by the presence of miR403 (dicot specific miRNA) in stevia. Another interesting observation was identification of miR894 homolog in stevia. miR894 was reported from *Physcomitrella* (Fattash et al., 2007), and in a few dicots such as sorghum (Zhang et al., 2011) tomato (Moxon et al., 2008) , cotton (Kwak et al., 2009) and cucurbits (Jagadeeswaran et al., 2012). Identification of miR894 in stevia and its absence in several other dicots raises several interesting evolutionary questions including whether homologs of miR894 were selectively lost in certain lineages.

As an important basal angiosperm species, stevia contains vast valuable information to understand the evolutionary history of the miRNAs from gymnosperms to monocots and eudicots. For example, miR1310 is reported as a specific miRNA in
gymnosperms plants (Morin et al., 2008), but it also identified in stevia based on our results and even present in *Linidendron chinese*. This implied that miR1310 is not a gymnosperm specific miRNA. It might disappear during the evolution of the angiosperms.

In stevia, five miRNA families (miR319, miR156/157 and miR165/166) have been found in more than forty plant species, and eight correspond to ancient miRNA already present in the common ancestor of embryophytes (miR156, miR159/319, miR160, miR165/166, miR171 and miR408), one in the common ancestor of tracheophytes (miR397 ) and six in the common ancestor of sporophytes (miR164, miR168, miR169, miR172, miR393, miR394 ) (Cuperus et al., 2004). The miRNAs sequenced in this study can definitely provide the information of *S. rebaudiana* miRNAs for further study on their gene regulation function, evolution and biogenesis.

Many conserved miRNA families have multiple loci and the members often varied by one or two nucleotides. Further, the largest miRNAs family size identified was miR166 that consisted of 17 members. miR156, miR159, miR167, miR319, miR396 and miR172 possessed 5, 8, 10, 8, 7 and 6 members respectively whereas other miRNA families such as miR157, miR160, miR169, miR858, miR894, miR2111 etc. had only one member detected in this library. The size of miRNAs families may be indicative of their function. Based on the results from the deep sequencing, different family members displayed drastically different expression levels. For example, the abundance of miR159 family varied from 6 reads to 11,759 reads in the deep sequencing. This was also the case for some other miRNA families, such as miR156 (from 3 read to 124 reads) miR167 (from 13 reads to 9,637 reads) and miR394 (from 2 reads to 1,554 reads). Abundance comparisons of different members in one miRNA family may provide useful information on the role that miRNAs play in the pre-flowering stage and even importance of the dominant member at that particular stage of plant growth.

### 5.4 Identification of novel miRNAs in *S. rebaudiana*

Computational approaches are successful in identifying novel or species specific miRNAs in many plants, but they require knowledge of the complete genome sequence, which is unavailable for most plant species. However, large genomic fragmented data in the form of GSSs, high-throughput genomics sequences (HTGSS), nonredundant nucleotides (NRs), as well as ESTs, are available for several plant species and can be used for identification of novel miRNAs. The NR database
contains finished genomic sequences and cDNAs. GSS and HTGS of GeneBank represent only short stretches of genomic sequence but can still provide a broader sampling of unfinished genomes (Sunkar and Jagadeeswaran 2008). All currently available miRNA discovery tools essentially require a reference/genomic sequence to identify a novel miRNA. Also, most of them are dependent upon identification of miRNA precursors. In such condition, non-availability of genomic sequences becomes a big limiting factor in the area of miRNA biology based research activities, including miRNA discovery. This has resulted into a sort of knowledge skew where most of the miRNAs have been reported only for those species whose genomic sequences are available or homologous sequences are known. Only 16% of total reported miRNAs in miRBase are from species whose genome is not sequenced. Also, the majority of these 16% miRNAs has been identified using homology search and exhibit a very few species specific miRNAs (Jha and Shankar, 2013). Thus, barring a few model organisms, there is almost negligible miRNA information for most of the species. In our case, although the S. rebaudiana full genome sequence is not available, the large number of stevia expressed sequence tags and nucleotide sequences are an excellent source for precursor identification.

Based on BLASTn searches and hairpin structure prediction, we found potential precursors for all the 24 novel miRNAs. Lacking of genomic information might be the main reason that leads to the identification of only 24 novel miRNAs. Their mature sequences were aligned with all the miRNAs registered in the miRNA database miRBase (http://www.mirbase.org/) (Kozmara and Griffiths- Jones 2011) and PMRD (http://bioinformatics. cau.edu.cn/PMRD/) (Zhang et al., 2010), CSRDB (http://sundarlab. ucdavis.edu/ smrnas/) and non-redundant sequences in Genebank. There was no homolog matching with them, which implied the 24 miRNAs found in our data are novel and might be stevia-specific miRNAs and could play more species-specific roles.

Out of the total 70,081,886 hits to the database (length: 3,637,171) containing 5,646 EST and 29,874 nucleotide sequences (Total: 35520), sequences better than e-value 10 i.e. 3,186,710 sequences were taken into further consideration. Out of these sequences, sequences showing exact matches i.e. 2,401 sequences were filtered for further secondary structure prediction. A distinguishing feature of miRNAs is the ability of their pre miRNA sequences to adopt canonical stem loop hairpin structure. Our analysis of sRNA sequences in sRNA library has identified several potential novel miRNAs of 19 to 24 nt in length with characteristic fold-back structures of miRNA precursors (Meyers et al., 2008).
On the contrary 24 novel miRNAs were predicted according the criterion mentioned before. All the 24 followed secondary structure criteria forming a stem loop structure with miRNA sequence sitting at one arm of the hairpin and negative minimal folding energy. The A+U contents of predicted miRNAs were found in the range of 42.1 -75% . Out of these 24 sRNAs, sequences of Stv_2, Stv_3, Stv_6, Stv_7, Stv_10, Stv_13, Stv_14, Stv_16, Stv_17 and Stv_19 start with 5'U, the characteristic feature of miRNAs. Though stv_7, stv_10 and stv_13 contained less than 50% A+U content but they start with U. Keeping this in view, these three sRNAs could be considered as putative stevia miRNAs. Earlier accepted findings state that non-conserved miRNAs are usually expressed at lower levels showing a tissue or development specific pattern (Song et al., 2010) as was seen in case of stevia where the read number of novel miRNAs was much lower than that for conserved miRNAs. Majority of them had only 1–5 sequenced clones. This suggests miRNAs identified in stevia might represent only a meager portion of novel miRNAs due to the fact that sRNA library was constructed from young plant leaves under normal conditions. In general, novel miRNAs represent either lineage-specific or species specific miRNAs and are expressed at low levels.

Further to experimentally confirm the presence of first 12 out of above 24 predicted novel miRNAs sRNA fraction, including miRNAs, was polyadenylated and reverse-transcribed with a poly(T) adapter into cDNAs for PCR using the miRNA-specific forward primer and the sequence complementary to the poly(T) adapter as the reverse primer. Forward primers were designed based on mature miRNA sequences (Table 3.2). If the Tm of a mature miRNA was <60°C, it was adjusted by adding Gs or Cs to the 5’ end and/or As to the 3’ end of the miRNA sequence (Shi and Chiang 2005). A primer corresponding to a 20-bp segment at the 3’ end of a *Stevia rebaudiana* 5.8S rRNA gene was used as a reference control. To further confirm the specificity of the amplification, we analyzed PCR samples on 2.5% agarose gels with EtBr visualization of bands. Fragments were gel-purified, cloned into the pGEM-T Easy vector (Promega) and sequenced.

Out of 12, nine miRNAs i.e. stv_1, stv_2, stv_3, stv_4, stv_5, stv_7, stv_9, stv_11 and stv_12 showed positive amplification confirming their presence and the fact that most of the miRNAs predicted through the computational workflow are true and are present in our model plant i.e. *S. rebaudiana*. Further, as seen in the agarose gel images no amplification was seen for only three novel miRNAs out of 12 i.e. stv_6, stv_8 and stv_10 which indicated that they may have very low expression at the time of sample collection.
As an evolutionary important plant genus, stevia might contain a certain amount of novel genes, including miRNA genes. With the availability of more genomic sequence information on stevia genus, more miRNAs will be identified in this species. The cloning and identification of these genes and figuring out their regulation relationships would be very helpful for exploiting new genes and regulatory pathways and their evolution in plant.

5.5 Target Prediction

To assess and define a putative function for a miRNA in plant, a further step of target identification is necessary. Currently, the most efficient tool available for this is the bioinformatics approach facilitated by the high degree of homology between miRNA and its target sequences in plants (Song et al., 2010). Analysis of several targets has now confirmed this prediction, making it feasible to identify plant miRNA targets (Llave et al., 2002, Park et al., 2002, Reinhart et al., 2002). Based on their functional analysis, miRNAs appear to be involved in almost all aspects of plant growth and development (Jones-Rhoades et al., 2006), as well as biotic and abiotic stress responses (Katiyar-Agarwal and Jin, 2010; Sunkar, 2010; Sunkar et al., 2007).

Candidate targets of the stevia miRNAs were predicted through BLAST, and further confirmed them by alignment with their orthologs in Arabidopsis. Our analysis reveals that most of the predicted targets in stevia have a conserved function with miRNA targets in Arabidopsis and these miRNA target sequences are also highly conserved among a wide variety of plant species as reported by Floyd and Bowman (Floyd and Bowman 2004). Even though miRNAs generally function as negative regulators of gene expression by mediating the cleavage of target mRNAs (Llave et al., 2002) or by repressing their translation (Chen, 2004), the cleavage of target mRNAs appears to be the predominant mode of gene regulation by plant miRNAs as was observed in case of stevia (Sunkar et al., 2006).

5.5.1 A. thaliana as reference genome for target prediction

As usually the targets are conserved between plant species we predicted putative target genes for 27 out of 34 conserved families and for 17 out of 23 stevia specific or novel miRNAs using Arabidopsis thaliana as reference genome through psRNA Target program.

Consistent with previous reports, most of these targets in stevia were plant-specific transcription factors, such as AP2, NAC, SBP and the ARF family. Our
results showed that *S. rebaudiana* miRNA targets encoded not only indispensable transcription factors, but also non-transcriptional factor proteins involving in diverse physiological processes. For example, HD-ZIP family transcription factor, was predicted to be target of conserved miR166 family in stevia. Similarly in *Medicago truncatula* and *P. densata* HD-ZIP family transcription factor was found to be target of miR166 which are important in their root and nodule development (Boualem et al., 2008). Further, HD-ZIP proteins also regulates vascular development as well as lateral organ polarity and meristem formation as seen in *Arabidopsis* where ATHB15, a member of the HD-ZIP family, is predominantly expressed in vascular tissues, suggesting that it may play some roles in plant vascular development in stevia too (Ohashi-Ito and Fukuda, 2003; Kim et al., 2005). Overexpression of miR166a results in decreasing ATHB15 mRNA levels and causes accelerated vascular cell differentiation from cambial/procambial cells and consequently produces an altered vascular system with expanded xylem tissue and an interfascicular region (Kim et al., 2005). This regulation mechanism may exist in all vascular plant species including our model plant *S. rebaudiana* (Floyd and Bowman, 2004; Kim et al., 2005). The importance of miR166 family at this preflowering stage of tissue collection was confirmed as the miRNA166 family was not only abundant during this stage of sample collection, but also had more family members than other miRNA families.

The miR165 and miR166 miRNAs are very similar in sequence and derive from multiple loci in the genome (Reinhart et al., 2002). They have been determined to regulate organ polarity by cleavage mRNAs (Mallory et al., 2004) or transcription repression of the class III homeodomain leucine zipper (HD-zip) family genes (Bao et al., 2004). Polarity of lateral organs is established through the antagonistic interactions between two major groups of genes, the HD-zip family such as PHABULOSA (PHB) and PHAVOLUTA (PHV) which are specific targets predicted in our case. So, this confirms the fact that miR165/166-mediated regulation of HD-zip genes is evolutionarily conserved among various plant species.

Further, AGO1 is one of the essential components involved in the miRNA biogenesis. The prediction of AGO1 to be the target of stv-miR168 and AGO2 of stv-miR403 implies that the miRNA biogenesis process is self-regulated in *S. rebaudiana*. Previously reported, overexpression of a miR166-resistant version of AGO1 appears to affect miRNA function because miRNA targets overaccumulate and the plants show phenotypes similar to those of miRNA biogenesis mutants such as dcl1, hen1 and hyl1 (Vaucheret et al., 2004). The regulation of genes involved in
sRNA metabolism or function by miRNAs is perhaps a feedback mechanism to ensure a certain level of activity of sRNA pathways.

It was found that target sites for conserved miRNAs in this plant were similar or functionally related to validated plant miRNA targets e.g. most members of the Squamosa Promoter Binding Protein Like (SPL) transcription factor family are targeted by miR156 in plants (Rhoades et al., 2002). For instance, miR156 targets 11 of the 17 SPL genes in Arabidopsis and SPLs affect diverse developmental processes such as leaf development, shoot maturation, phase change and flowering in plants. Similarly, in stevia miR156 has been predicted to target SPL. Additionally, miR156 has been shown to be involved in floral development and phase change by targeting members of squamosa promoter binding protein like (SPL) plant-specific transcription factors. Recent results indicated that overexpression of miR156 affects phase transition from vegetative growth to reproductive growth, including the quickly initiation of rosette leaves, a severe decrease in apical dominance, and a moderate delay in flowering (Schwab et al., 2005).

Further, in Arabidopsis, miR159 has been shown to be involved in the regulation of seed dormancy and germination by targeting MYB101 as in our case, positive regulator of ABA responses during germination. ABA is a key regulator of seed maturation and dormancy (Finkelstein et al., 2008). Many ABA signal transduction proteins are involved in seed development and germination (Finkelstein et al., 2008, Nakashima et al., 2009, Yamagishi et al., 2009, Kinoshita et al., 2010).

The phytohormone auxin plays critical roles during plant growth and development, including stem elongation, phototropic and gravitropic responses, apical dominance, lateral and adventitious root formation. The plant response to auxin involves the short-lived auxin/indole-3-acetic acid (Aux/IAA) proteins, the ARFs, and the components of the protein degradation pathway (Dharmasiri and Estelle, 2004). Recently, a number of genes in auxin signaling are confirmed or predicted as targets of miRNAs. Similarly in stevia, ARFs are a class of targets of miR160 family as is seen in other plants. ARFs are important components of auxin signal transduction (Guilfoyle and Hagen, 2007). Expression of a miR160-resistant version of ARF17 (5mARF17) leads to pleiotropic developmental abnormalities, such as leaf serration, leaf curling, early flowering, altered floral morphology, and reduced fertility (Mallory et al., 2005). This indicates that miR160-mediated regulation of ARF17 is critical for various aspects of development of stevia.
In stevia NAC1, which encodes a transcription factor acting downstream of TIR1 to promote lateral root formation, is a target of miR164. miR164 guides the cleavage of NAC1 mRNA in vivo (Guo et al., 2005). In a recent study, T-DNA insertions in two of the three members of the miR164 family (mir164a, mir164b) lead to 1/4-1/3 wild type levels of total miR164 and cause an increase in NAC1 mRNA levels and a corresponding increase in lateral root number (Guo et al., 2005).

As seen in Arabidopsis miR319 in stevia targets TCP family transcription factor 4. In Arabidopsis, overexpression of miR319/Jaw leads to crinkly leaves and reduced levels of five TCP genes containing miR319/Jaw-binding sites (Palatnik et al., 2003). Further, overexpression of a miR319/Jaw-resistant form of TCP2 fully restores the crinkly leaf phenotype of miR319/Jaw overexpression (Palatnik et al., 2003).

Plants are sessile, and thus, have to cope with adverse environments. They have evolved sophisticated mechanisms to adapt to environmental stresses. At present, many miRNAs have been predicted and some have been confirmed experimentally to be involved in a variety of stress responses, including abiotic (Sunkar and Zhu, 2004, Zang et al., 2005) and biotic stresses (Zang et al., 2005), oxidative stress (Sunkar et al., 2006), mechanical stress in tree species (Lu et al., 2005), nutrient stresses (Fujii et al., 2005) and other environmental stimuli. These primary works will provide a foundation for future researches in the field like our case. Another validated mRNA targeted by miR408 codes a plastocyanin-like protein and is believed to mediate lignin polymerization. These results show that plant miRNAs can be induced by mechanical stress and may function in one of the most critical defense systems for structural and mechanical fitness.

Based on their functional analysis, novel miRNAs also appear to be involved in almost all aspects of plant growth and development (Jones-Rhoades et al., 2006), as well as biotic and abiotic stress responses using A. thaliana as reference genome (Katiyar-Agarwal and Jin, 2010, Sunkar, 2010, Sunkar et al., 2007). Stv_1 is predicted to target transcription factors including ethylene-responsive family protein and AT-GTL1 and AT-GTL2 transcription factors and several other interesting targets including retinoblastoma protein Rb which is considered to be a key regulator of G1/S phase transition by blocking S phase entry and cell growth and RING finger proteins are involved in numerous cellular processes including transcription, signal transduction, and recombination. Functions attributed to the RING domain itself
include protein–protein interaction and ubiquitination (Borden and Freemont, 1996, Lorick et al., 1999).

Further stv_2 targets both structural and functional components in plants including acetyl-CoA carboxylase involved in fatty acid synthesis and oxidoreductases involved in electron transfer. Stv_3 is predicted to target interesting components including ataxia telangiectasia mutated (ATM) which is a serine/threonine protein kinase that is recruited and activated by DNA double-strand breaks which phosphorylates several key proteins that initiate activation of the DNA damage checkpoint, leading to cell cycle arrest, DNA repair or apoptosis and ubiquitin thiolesterase which aids synthesis of ubiquitin, as the addition of ubiquitin can affect proteins in many ways: it can signal for their degradation via the proteasome, alter their cellular location, affect their activity, and promote or prevent protein interactions (Glickman and Ciechanover, 2002; Mukhopadhyay and Reizman, 2007; Schnell and Hike, 2003).

Further, Stv_4 targets include CYCP3;2 which is involved in diverse physiological responses including regulation of cell cycle, regulation of cyclin-dependent protein serine/threonine kinase activity and stomatal lineage progression and another interesting target includes TIR-NBS-LRR disease resistance protein The majority of disease resistance genes in plants encode these proteins and their precise role in recognition is unknown; however, they are thought to monitor the status of plant proteins that are targeted by pathogen effectors. Further it targets two important factors in the transcription machinery i.e. ANAC023 (Arabidopsis NAC domain containing protein 23) and transcription factor TGA6. Stv_5 and Stv_7 targets transposable element gene or retrotransposon which is a DNA sequence that can change its position within the genome, sometimes creating or reversing mutations and altering the cell's genome size. Stv_6 targets Late Embryogenesis Abundant proteins (LEA proteins) which are proteins in plants that protect other proteins from aggregation due to desiccation or osmotic stresses associated with low temperature (Goyal et al., 2005) and integral membrane transporter family protein (IMPs) which include transporters, linkers, channels, receptors, enzymes, structural membrane-anchoring domains, proteins involved in accumulation and transduction of energy, and proteins responsible for cell adhesion. Stv_8 targets pectinesterase which functions primarily by altering the localised pH of the cell wall resulting in alterations in cell wall integrity and WD-40 repeat family protein which are implicated in a variety of functions ranging from signal transduction and transcription regulation to cell cycle control, autophagy and apoptosis (Stirnimann et al., 2010).
Stv_9 targets lectin receptor kinases which are thought to play crucial roles during development and in the adaptive response to various stresses (Singh and Zimmerli, 2013) and a gametogenesis related protein. On the other hand, Stv_11, Stv_12 and Stv_13 possess single targets i.e. EDA23 protein tyrosine kinase which is involved in embryo sac development, protein amino acid phosphorylation, transmembrane receptor protein tyrosine kinase signaling pathway; ARID/BRIGHT DNA-binding domain-containing protein involved in regulation of transcription and lastly serine carboxypeptidase with active site serine involved in diverse functions ranging from catabolism to protein maturation respectively. Additionally, Stv_14, Stv_17 and Stv_20 possess the same target i.e. WD-40 repeat family protein further proving the fact that same target can be targeted by more than one miRNA.

Several interesting components are targeted by other novel miRNAs including Stv_15 which targets CYCLIN D3 involved in many aspects of cell biology and notably in the regulation of the cell cycle and paired amphipathic helix (PAH2) super family protein involved in regulation of transcription; Stv_17 targets WRKY transcription factors which are one of the largest families of transcriptional regulators in plants and form integral parts of signaling webs that modulate many plant processes (Rushton et al., 2010) and permease which are membrane transport proteins, a class of multipass transmembrane proteins that facilitate the diffusion of a specific molecule in or out of the cell by passive transport; Stv_19 targets HEAT repeat-containing protein involved in ribosome biogenesis and rRNA processing, choline monooxygenase and UDP-glucosyltransferase involved in various metabolic processes; Stv_21 targets pentatricopeptide repeat (PPR) super family which interact (often essentially) with mitochondria and other organelles (Lurin et al., 2004) and that they are possibly involved in RNA editing (Kotera et al., 2005) and cyclophilin-like peptidyl-prolyl cis-trans isomerise involved in protein folding; Stv_22 targets WRKY DNA-binding protein which are thought to play an important role in plant defense responses, plant hormone signaling, secondary metabolism and plant responses to abiotic stresses (Cheng et al., 2012). Lastly, Stv_24 is predicted to target endomembrane protein 70 protein family which are localized in the secretory pathway of eukaryotic cells and are involved in cell adhesion and phagocytosis (Benghezal et al., 2003).
The high degree of sequence complementarity between plant miRNAs and their target mRNAs has facilitated the bioinformatic prediction of miRNA targets. Plant miRNAs have been predicted or confirmed to regulate a wider variety of developmental and physiological processes than animal miRNAs (Dugas and Bartel, 2004). With these increasing evidence, it is concluded that miRNAs regulatory impact on plants is more pervasive than was previously suspected. miRNA target prediction in plants is easier due to high and significant complementarities of miRNA-mRNA targets (Lee and Ambros, 2001) and targets of most plant miRNAs possess a single perfect or near perfect complementary site in coding region (Adai et al., 2005). Assuming this to be generally the case, the Stevia transcript library (EST and nucleotide databases) was searched for complementarity with the sequences of identified conserved and novel miRNAs using the two target prediction web servers psRNATarget (http://bioinfo3.noble.org/psRNATarget/) and TAPIR (http://bioinformatics.psb.ugent.be/webtools/tapir). Almost all the targets predicted through psRNA target server were further authenticated by TAPIR which showed the same targets. Further, TAPIR and psRNA target program individually predicted additional putative targets. A deep insight into miRNA targets helps us in understanding the range of sRNA expression, regulation and their functional importance. To analyze probable sRNA targets is significant in plants because the complimentary sites of potential sRNA can exist anywhere along the target mRNA rather than at 3'UTR in case of animals (Zhang et al., 2006). As usually the targets are conserved between plant species but to know specific targets in our model plant S. rebaudiana we predicted putative target genes for 10 out of 34 conserved families and for 13 out of 23 stevia specific or novel miRNAs taking stevia transcript library as reference library. Similarly, mode of action of miRNAs on targets was found to be predominantly through cleavage which is usually the case seen in plants as observed above taking A. thaliana as reference genome (Llave et al., 2002, Jones Rhoades and Bartel, 2004).

Target prediction was done for conserved miRNAs using Stevia transcript library to confirm the conserved nature of targets between plant species and to know more relevant targets in our model plant itself i.e. S. rebaudiana. Parallel targets were observed for miR156, miR160, miR166, miR167 and miR319, miR393, miR396, miR414, miR856 which were Glucan endo-1,3-beta-glucosidase, early nodulin-like protein, catalase, light harvesting chlorophyll a/b binding protein precursor, peroxidase, RING/U-box domain-containing protein and elongation factor 200.
respectively (Rhoades et al., 2002) as seen in other plant species including Arabidopsis.

Further, miR414 is predicted to regulate multiple targets including RING/U-box domain-containing protein, phytochrome E, WD-repeat protein, vacuolar ATP synthase subunit E, SIN3 component, translation elongation factor, molecular chaperone Hsp90, ALF domain class transcription factor, programmed cell death 2 C-terminal domain-containing protein, resistance protein, NAC, DNA-damage repair protein and several proteins involved in synthesis i.e. 60S acidic ribosomal protein etc. Further, miR414 is predicted to have interesting stevia genes as targets like kaurene synthase (KS22-1), calmodulin and UDP-glycosyltransferase 73E1. Kaurene synthase is a vital late gene involved in steviol glycoside biosynthesis pathway. Further targets were predicted for certain more conserved miRNAs including miR166, miR167, miR319, miR 396 and miR408, miR856 and miR1310. miRNAs regulate gene expression predominantly by cleavage (Table 4.8) due to high complementarity of miRNA and targets.

Putative targets were also predicted for some novel miRNAs. Stv_1 targets one of the chaperonin_like superfamily protein whose common function is to sequester non-native proteins inside their central cavity and promote folding by using energy derived from ATP hydrolysis. On the other hand, Stv_2 targets phosphoglycerate kinase (PGK) which is a monomeric enzyme that catalyzes the transfer of the high-energy phosphate group of 1,3-bisphosphoglycerate to ADP, forming ATP and 3-phosphoglycerate. This reaction represents the first of the two substrate-level phosphorylation events in the glycolytic pathway. Stv_7 targets pectinacetylesterase family protein which is located in cell wall or membrane and exhibits carboxylesterase activity. Further, Stv_10 is predicted to target ribulose-1,5-bisphosphate carboxylase activase (P-loop NTPases) involved in diverse cellular functions as well as lupus la ribonucleoprotein or RRM (RNA recognition motif) found in proteins involved in post-transcriptional gene expression processes including mRNA and rRNA processing, RNA export, and RNA stability. Additionally, we predicted a few genes with unknown function and hypothetical genes for miRNA targeting e.g. Stv_8.

Stv_14 targets 3-hydroxyisobutyryl-CoA hydrolase which participates in 3 metabolic pathways: valine, leucine and isoleucine degradation, beta-alanine metabolism, and propanoate metabolism and phytochrome b involved in photoperiodic induction of flowering, chloroplast development, leaf senescence and
leaf abscission. Stv_16 targets Zinc finger A20 and AN1 domain-containing stress-associated protein involved in environmental stress responses induced by cold, dehydration and salt stresses. Further, Stv_17 targets some interesting genes i.e. Histone H3 which is one of the five main histone proteins involved in the structure of chromatin in eukaryotic cells (Bhasin et al., 2006), 26S proteasome non-ATPase regulatory subunit that acts as a regulatory subunit of the 26S proteasome which is involved in the ATP-dependent degradation of ubiquitinated proteins and armadillo repeat-containing kinesin-like protein 2 which is involved in the control of epidermal-cell morphogenesis in roots and helical growth of roots by promoting microtubule depolymerization and limiting the accumulation of endoplasmic microtubules (Sakai et al., 2006). Further TMV resistance protein is targeted by more than one miRNA i.e. stv_17 and stv_19 and other stress protein i.e. hypersensitive induced protein is targeted by stv_18 further, proving the indispensable roles of miRNAs in various stress responses.

Further as miRNAs regulation is seen at both transcriptional and translational level, stv_21 is found to target GATA transcription factors which are a family of transcription factors characterized by their ability to bind to the DNA sequence “GATA” (Ko and Engel, 1993) and Stv_22 on the other hand targets Elongation factor P, an important component of translation machinery and DEAD box proteins are considered to be RNA helicases and many have been found to be required in cellular processes such as pre-mRNA processing and rearranging of ribonucleoproteins (RNP) complexes (Staley and Guthrie, 1998).

No targets were predicted for stv_10, stv_11, stv_12 and stv_23 because targets predicted using A. thaliana were insignificant as they bound in the non-coding regions of the gene whereas no targets were predicted for them using S. rebaudiana transcript library as reference. Target of only 13 out of 23 novel miRNAs were predicted speculating that rest of the miRNA may be typically present in stevia and needs further experimental work to locate their gene targets. The above described information regarding predicted targets could be utilized efficiently to assess the regulatory roles of these novel miRNAs in stevia. Further if any of these putative novel miRNA is found to have control over the regulation of the early and late genes in steviol glycosides biosynthetic pathway as seen in case of conserved miRNA 414 can help to manipulate the diterpenoid contents.

The tissue- or cell-type-specific functions of the miRNAs discussed above suggest that miRNA gene expression is precisely regulated in the plant. It is
conceivable that the abundance of a certain miRNA in cells can be regulated at multiple levels, such as the transcription of the gene, the processing of the primiRNA by DCL1, the methylation of the miRNA by HEN1, the loading of the miRNA into RISC, and the export of the miRNA into the cytoplasm, and even the potential transport of the miRNA in or out of cells or tissues (Yoo et al., 2004). Although how miRNAs themselves are regulated is currently unknown, available data show that miRNA accumulation exhibits tissue or cell type specificity and responses to stimuli.

FUTURE PERSPECTIVES

➢ Our study lays a strong foundation for experimental validation of these predicted miRNA-target interactions in S. rebaudiana using several methods such as RACE, reporter assays, microarray and proteome analyzes which will help in understanding the complex regulatory networks between miRNA, secondary metabolism and various abiotic and biotic stress responses.

➢ Finally, a better understanding on the miRNA regulation mechanisms in plants will make it possible to design artificial miRNAs that may be used as efficient tools for controlling gene expression at will (Schwab et al., 2006; Alvarez et al., 2006).

➢ Further if any of the predicted novel miRNAs is confirmed to have control over the regulation of the early and late genes in steviol glycosides biosynthetic pathway we can manipulate the diterpenoid contents, hence, alter the sweetness indices of S. rebaudiana.

➢ In future, with the availability of genome sequence of S. rebaudiana, we can predict more miRNA and their respective coding genes and targets of miRNA in non-coding regions of the genome such as introns and UTRs.