Suresh Kumar Gahlawat · Raj Kumar Salar
Priyanka Siwach · Joginder Singh Duhan
Suresh Kumar · Pawan Kaur *Editors*

Plant Biotechnology: Recent Advancements and Developments

Springer

Plant Biotechnology: Recent Advancements and Developments

Suresh Kumar Gahlawat • Raj Kumar Salar
Priyanka Siwach • Joginder Singh Duhan
Suresh Kumar • Pawan Kaur
Editors

Plant Biotechnology: Recent Advancements and Developments

Springer

Editors
Suresh Kumar Gahlawat
Department of Biotechnology
Chaudhary Devi Lal University
Sirsa, Haryana, India

Priyanka Siwach
Department of Biotechnology
Chaudhary Devi Lal University
Sirsa, Haryana, India

Suresh Kumar
Department of Biotechnology
Chaudhary Devi Lal University
Sirsa, Haryana, India

Raj Kumar Salar
Department of Biotechnology
Chaudhary Devi Lal University
Sirsa, Haryana, India

Joginder Singh Duhan
Department of Biotechnology
Chaudhary Devi Lal University
Sirsa, Haryana, India

Pawan Kaur
Department of Biotechnology
Chaudhary Devi Lal University
Sirsa, Haryana, India

ISBN 978-981-10-4731-2 ISBN 978-981-10-4732-9 (eBook)
DOI 10.1007/978-981-10-4732-9

Library of Congress Control Number: 2017942816

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

The present century is a century of knowledge, indeed a century of the mind, and it belongs to biotechnology, which has brought unprecedented advances in all walks of life. The origin of biotechnology can be traced back to prehistoric times, when microorganisms were used for processes like fermentation. The advent of biotechnology has opened up a wide horizon in the field of biological research. It has developed into an important force in the creation of employment, production, and trade, and affects national economy and the quality of human life throughout the world. The range and the significance of biotechnological activities are expanding at an ever-increasing rate so that the impact of biotechnology on the world economy and human welfare is likely to be far greater than that of any other discipline. Our effort is to introduce students and researchers to cutting-edge techniques and applications of biotechnology. Therefore, chapters were selected on climate change, global warming, genetic engineering, crop improvements, RNAi technology and nanoparticles, etc., for this book.

The book entitled *Plant Biotechnology: Recent Advancements and Developments* will be highly useful for students, teachers, and researchers in all disciplines of life sciences, agricultural sciences, and biotechnology in universities, research institutions, and biotechnology companies. It will provide the readers a comprehensive knowledge of topics on genomics, plant biotechnology, and bionanotechnology. The chapters have been written with special reference to the latest developments in the above broader areas of biotechnology that impact the biotechnology industry. A list of references at the end of each chapter is provided for the readers to learn more about a particular topic. Typically, these references include basic research, research papers, review articles, and articles from the popular literature.

We as editors are grateful to the contributing authors for providing necessary expertise and commitment. Without their contribution, this book would not have been possible. Our sincere thanks go to the University Grants Commission, New Delhi, for sponsoring the National Conference on Biotechnology: Emerging Trends, which was organized by the Department of Biotechnology, Chaudhary Devi Lal University, Sirsa, India, during February 2016 and attended by most contributing

authors. Last but not the least, we also thank our family members for their constant support and encouragement while we were working on this project.

Sirsa, Haryana, India

Suresh Kumar Gahlawat
Raj Kumar Salar
Priyanka Siwach
Joginder Singh Duhan
Suresh Kumar
Pawan Kaur

Contents

About the Editors

Suresh Kumar Gahlawat, Ph.D. is the chairperson, Department of Biotechnology, and dean, Faculty of Life Sciences, Chaudhary Devi Lal University (CDLU), Sirsa, India. He received postdoctoral BOYSCAST fellowship and DBT Overseas Associateship from the Ministry of Science & Technology, Government of India, for carrying out research at FRS Marine Laboratory, Aberdeen, the UK. He has completed four R&D projects from UGC, ICAR, and Government of Haryana. His research interests include the development of molecular diagnostic methods for bacterial and viral diseases. He published more than 70 research papers in journals of national and international repute, written five books, and supervised M.Phil. and Ph.D. research work of 12 students.

Raj Kumar Salar, Ph.D. is a professor at the Department of Biotechnology, CDLU, Sirsa, India. He was awarded a postdoctoral fellowship from the Ministry of Education, Slovakia, to pursue postdoctoral research at Slovak University of Technology, Bratislava, and has also visited Japan, Norway, Austria, and Hungary for academic pursuits. Dr. Salar received several R&D projects from UGC, New Delhi, HSCST, Chandigarh, and DST, New Delhi. He was felicitated with King Abdulaziz City for Science and Technology (KACST) award for the best paper published in 3 Biotech. He has supervised several Ph.D. and M.Phil. students for their research. Dr. Salar is also a reviewer of several international journals.

Priyanka Siwach, Ph.D. is a professor at the Department of Biotechnology, CDLU, Sirsa. She was awarded with Silver Jubilee Gold Medal for her Ph.D. thesis by CCS Haryana Agricultural University, Hisar. She has received R&D projects from UGC, New Delhi. Her current research areas include genome annotations, identification and functional annotation of miRNAs, and molecular marker analysis for improvement of quality traits. She has also worked on micropropagation and biochemical characterization of medicinal plants and filed two patents in this field. She is a reviewer of several national and international journals.

Joginder Singh Duhan, Ph.D. (Microbiology) is a senior assistant professor at the Department of Biotechnology, CDLU, Sirsa. He is the recipient of Young Scientist Award given by the Association of Microbiologists of India and K.K. Nanda award for best research and several other best paper awards. He has more than 70 research papers and edited two books published by Springer and Narosa. He has received R&D projects from UGC and CSIR, New Delhi. He has supervised the research work of 15 M.Phil. and 5 Ph.D. students. His specialization includes microbial biotechnology and fermentation technology. He is a reviewer of several national and international journals.

Suresh Kumar, Ph.D. is presently working in the Department of Biotechnology, CDLU, Sirsa. His specialization includes microbial and environmental biotechnology. He has published more than 15 research papers in journals of national and international repute.

Pawan Kaur, Ph.D. is presently working in the Department of Biotechnology, CDLU, Sirsa. She has post-Ph.D. research experience in the area of drug delivery. Her area of interests includes synthesis of nanoparticles, nanocomposites, and antimicrobial activity and cytotoxicity of nanoformulations in biotechnology. She has published more than 15 research papers in journals of national and international repute.

Advances in Computational Tools for Plant microRNA Identification

1

Megha Sihag, Khushboo Sethi, S.K. Gahlawat, and Priyanka Siwach

Abstract

Continuous development in biotechnological techniques has led to major breakthrough discoveries in life sciences, one of which is the finding of a particular class of small RNA molecules known as microRNAs (miRNA). miRNAs have been found to regulate various biological activities in all life forms. Though plant miRNAs were discovered later than animal miRNAs, these have been found to reveal remarkable importance in gene regulation during plant development as well as toward responding to any stimuli. Unraveling the entire mystery of these small molecules is the first step to gaining a better understanding of their function. Initially, miRNAs were identified largely by experimental techniques like forward genetic screening, cloning, and microarray, but these techniques were expensive, more time-consuming, and unable to reveal poorly expressed RNA molecules. Advances in computational tools, made during the last decade, have done wonders in miRNA research work. Many novel miRNAs were discovered using these approaches in plant systems, though confirmation requires experimental validation. In this chapter, an effort has been made to understand the characteristic features of plant miRNA compared to animal miRNAs and to review the recent advances in various computational tools made for identification of plant miRNAs.

M. Sihag • K. Sethi • S.K. Gahlawat • P. Siwach (✉)
Department of Biotechnology, Chaudhary Devi Lal University, Sirsa 125 055, Haryana, India
e-mail: psiwach29@gmail.com

S.K. Gahlawat et al. (eds.), *Plant Biotechnology: Recent Advancements and Developments*, DOI 10.1007/978-981-10-4732-9_1

1.1 Introduction

MicroRNAs (miRNA) are small noncoding RNAs that play an important role in regulation of gene expression at posttranscriptional stage (Ambros 2001; Carrington and Ambros 2003; Bartel 2004). The first miRNA (lin-4) was discovered by Ambros and colleagues in 1993 while working on developmental regulation in *C. elegans* (Lee et al. 1993). For seven years, lin-4 was considered as an unusual thing, until the discovery of a second miRNA in *C. elegans*, called let-7 (Reinhart et al. 2000). Following this, let-7 homologs were identified in many vertebrate species, including humans (Pasaquinelli et al. 2000), and were later on found to be evolutionarily conserved molecules. The potential of miRNA in regulation of gene expression was not recognized until 2001 (Ambros 2001; Lau et al. 2001; Lagos-Quintana et al. 2001). Extensive follow-up work by many laboratories revealed the concepts of miRNA biogenesis and function (Kurihara and Watanabe 2004; Lee et al. 2002; Bartel 2004; Yu et al. 2005).

Plant miRNAs were identified nearly a decade later than animal miRNAs (Llave et al. 2002; Reinhart et al. 2002; Park et al. 2002). Since then, different approaches have been undertaken to identify new miRNAs. A public repository for all published miRNA sequences and associated annotation was established in 2002, named as MicroRNA Registry, initially, and as miRbase, later on. The first release of miRBase in 2002 contained 218 miRNA loci from five species, while the latest miRbase release (v21, June 2014) contains 28,645 entries from 206 species, of which 6150 miRNAs are from different plant species as on June 2016, indicating the intensity and speed of miRNA research projects. A separate database for plant miRNA, plant microRNA database (PMRD) has also been developed to integrate the data pertaining to plant miRNAs available from public resources (Zhang et al. 2010). The database was established in 2008 and has been updated regularly. One update was made on February 14, 2014, with a total of 10,594 miRNA entries from 128 plant species in the PMRD (http://bioinformatics.cau.edu.cn/PMRD). In November 2014, PMRD was changed to PNRD (Plant Noncoding RNA Database), to help biologists with their research in the RNA world. Noncoding RNA (ncRNA) is a functional RNA molecule that is not translated into a protein and is highly abundant and functionally important. The latest update of PNRD was made on May 2, 2016, with a total of 28,214 ncRNA entries from 166 plant species in the PNRD in which 16,390 were miRNA entries from 154 species. (http://structuralbiology.cau.edu.cn/PNRD/). Some other plant-based databases are MicroPC (http://www.biotec.or.th/isl/micropc) and PmiRKB (http://bis.zju.edu.cn/pmirkb/), which are basically annotation databases and have greatly enhanced plant miRNA research. Since the identification of first miRNA in plants, a lot of advancement has taken place in the miRNA identification approach and different tools have been developed for this. Various efforts have been made from time to time to discuss and compare various strategies adopted for miRNA identification (Meyers et al. 2008; Unver et al. 2009; Mendes et al. 2009; Hu et al. 2012). Since the computational tools and laboratory techniques, as well as the understanding of miRNA genes, are continuously evolving, this chapter is another effort in this direction to review various advances made for plant miRNA identification.

1.2 Characteristic Features of Plant miRNAs Versus Animal miRNAs

In both animal and plant cells, miRNA genes are transcribed by polymerase II to capped and polyadenylated primary miRNA (pri-miRNA) transcript (Lee et al. 2004). These pri-miRNAs are processed to form stem-loop precursor miRNA (pre-miRNA) (Kurihara and Watanabe 2004), with a minimal folding free energy index (MFEI) which is further processed to mature miRNAs. The latter enter a ribonucleoprotein complex known as the RNA-induced silencing complex (RISC), which regulates targeted gene expression (Bartel 2004). However, the biogenesis of plant and animal miRNAs differ in some aspects (Fig. 1.1). Studies have revealed larger and more variable precursor miRNA molecules for plants than for animals (Reinhart et al. 2002). In animals, processing of pri-miRNA to the pre-miRNA is mediated by

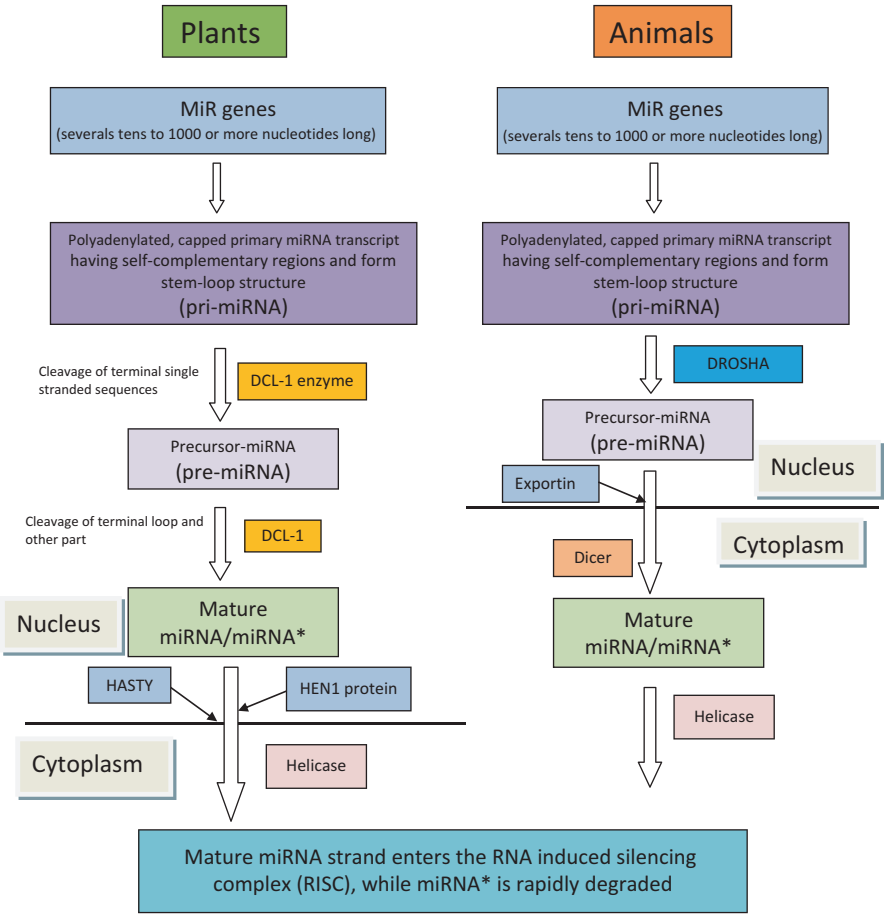

Fig. 1.1 Comparison of miRNA biogenesis mechanism in plants and animals

two enzymes Drosha and Dicer, in nucleus and cytoplasm respectively (Kim et al. 2009). On the other hand, in plants, a single enzyme (Dicer like-1) carries out both the processing steps in the nucleus only (Fahlgren et al. 2007), the same enzyme carrying out both cleavage steps, indicating that the stem-loop intermediate is a very transient structure in plants, compared to in animals. A stronger preference for a U at the 5′ terminus of the plant miRNAs has also been observed (Lau et al. 2001; Reinhart et al. 2002; Bartel and Bartel 2003). The differences between plant and animal miRNAs do not pertain to biogenesis only, but are also visible in their conservation, genomic organization, and mode of action (Millar and Waterhouse 2005; Zhou and Guan 2012).

Plant miRNAs are less conserved than animal miRNAs; usually only the mature miRNAs are conserved in plants instead of miRNA precursors that are usually conserved in animals (Bartel 2004). Phylogenetic studies indicate that plant miRNAs have a very deep origin in plant phylogeny, since the last common ancestor of bryophytes and seed plants (Zhang et al. 2005). The age of plant miRNA is suggested to be comparable to the age of miRNA regulation in metazoans; however, no evidence yet shows that plant miRNAs and animal miRNAs have a common ancestor. Both plant and animal miRNA genes are predominantly located in the intergenic regions, though significant numbers of animal miRNAs are also located in the introns of pre-mRNAs; for example, of the human miRNA genes, ~25% are encoded within introns (Bartel 2004). However, such observations are not common for plant miRNAs. Further, in plants, most miRNAs are encoded by their own primary transcript, with a few exceptional cases such as miR395 (which is present four times within a single transcript in rice), while miRNA clusters appear to have developed to a much greater extent in animals (Millar and Waterhouse 2005).

Both plant and animal miRNA, after entering RISC, bind to messenger RNA (mRNA) and inhibit gene expression. Substantial differences are seen in plant and animal miRNAs for target recognition and mode of gene regulation. Plant miRNA can bind to perfect or near to perfect complementary sites; most target mRNAs generally have a single complementary site in plants, and this site can be found anywhere along the target mRNA.

Plant miRNAs can silence the target gene through RNA degradation as well as translational repression (Brodersen et al. 2008). Some reports exist where plant miRNAs cause reduced level of proteins in plants but not mRNA (Cullen 2011). Animal miRNAs, on the other hand, usually bind to target mRNAs through imperfect complementarity at multiple sites located at the 3′ untranslated regions (UTR) and stop the ribosomal movements along the mRNA (Carrington and Ambros 2003).

From the above discussions, it is obvious that though plant and animal miRNAs have considerable similarities, they do have slight differences; these differences, if ignored during miRNA identification steps, will lead to generation of false-positive data.

In addition to above differences, extra cautious steps are needed for plant miRNA identification because plants have relatively large and complex small RNA populations within which miRNAs are often a minority. By contrast, most of the small RNAs in the animals are miRNAs (Meyers et al. 2008.). The complex small RNAs population in plants is largely due to the plant-specific RNA polymerase IV/RNA

polymerase V (PolIV/PolV)–dependent short interfering RNAs (siRNAs) as well as secondary siRNAs, some of which are trans-acting. The presence of these diverse, endogenous siRNA populations in plants makes the precise identification of miRNAs more challenging (Meyers et al. 2008).

1.3 Methods of Plant miRNA Identification

Presently a number of methods are available which have not only speeded up miRNA identification but have also enhanced the efficiency and accuracy. Although miRNAs were first discovered through forward genetic screens in roundworms (Lee et al. 1993; Reinhart et al. 2000), this method has not been successful for miRNA gene identification in plants. The first plant miRNA was discovered by cloning experiments (Llave et al. 2002; Reinhart et al. 2002). Intensive refinement in cloning techniques for small RNA molecules has helped to successfully identify many new plant miRNA (Lu et al. 2005). However, cloning is biased toward RNAs that are expressed highly and broadly. MicroRNAs expressed at low levels or only in specific cell types or in response to certain environmental stimuli are more difficult to clone. Sequence-based biases in cloning procedures might also cause certain miRNAs to be missed (Ambros et al. 2003; Mendes et al. 2009).

Other experimental approaches have also been used to investigate new miRNAs such as Northern blot, microarray, and in situ hybridization, but these are also tedious and time-consuming. The advent of next-generation sequencing (NGS) technology reduced the cost for discovery and offers significant advantage of identifying even the less abundant miRNAs. It also provides a more reliable and sensitive method to quantify known miRNAs (Table 1.1). However, to discover new miRNAs from NGS data, some kinds of miRNA prediction algorithms with proper computational infrastructure are required (Friedlander et al. 2008). Several computational tools have been developed to complement experimental approaches to identify and validate novel miRNAs from high-throughput platforms such as NGS (Huang et al. 2011).

The main miRNA characteristics used by different computational tools are their length, high sequence conservation among species, and structural features like hairpin and minimal folding free energy (Li et al. 2010). Presently, various computational plant miRNA identification methods exist, which can be broadly divided into two main strategies: comparative-genomics-based methods and ab initio methods (Table 1.2). However, more specific and further subclassification can also be applied.

Comparative-genomics-based approach is one of the most extensively developed methods for miRNA detection, both in plants and in animals (Bonnet et al. 2004; Jones-Rhoades and Bartel 2004). A comparative genomics study across taxa has shown that many mature miRNAs are evolutionarily conserved from species to species in plants, though pre-miRNAs are less conserved. Because the mature miRNA sequences are short (~22 nt), sequence alignment tools like BLAST (Altschul et al. 1990) can only find the nearly perfect matches due to the large number of irrelevant hits. Sequence alignment alone may fail to detect the distant homologs that diverge in sequence yet are conserved in structure. So, more sensitive

Table 1.1 Experimental tools applied for discovery of plant miRNA

Techniques used	Constraints	Some examples of work reference
Cloning	Time-consuming	Lu et al. (2005), Sunkar et al. (2005), Xie et al. (2005) and Zhao et al. (2012)
	Data is noisy (may clone and sequence a large number of degraded RNA fragments from the samples).	
	May not detect miRNAs that have low expression levels	
Microarray	Time-consuming	Chambers and Shuai (2009), Ding et al. (2011) and Liu et al. (2008)
	Low chances of detection of miRNA with low expression level	
In situ hybridization	Sensitive	Zhao et al. (2012)
	Time-consuming	
	Tedious	
Next-generation sequencing	To discover new miRNAs from NGS data, some kinds of miRNA prediction algorithms with proper computational infrastructure are required	He et al. (2013) and Jiao et al. (2011)

approaches that can consider both sequence and structure conservation are needed. On the other hand, ab initio methods predict miRNAs in a single genome without using conservation of structure or comparative sequence analysis. The number of nonconserved miRNAs is estimated to be relatively large, which enables the identification of completely novel miRNAs for which no close homologs are known. Unlike comparative genomics approaches, ab initio approaches can discover species-specific miRNAs without known homology sequences (Li et al. 2010). Ab initio methods, in general, do not depend entirely on known genetic sequence information but use it as a training set to extract features and then make use of a machine-learning classification algorithm to identify candidate clips. Previously, an attempt has been made to compare and discuss various ab initio methods for miRNA discovery in metazoans as well as plants (Wu et al. 2012). At present, many tools belong to each category. Some of the popular ones are discussed below.

1.3.1 microHARVESTER

This is a popular homology-based tool introduced for plant miRNA identification (Dezulian et al. 2006). It takes advantage of the conservation of mature miRNA. It uses a BLAST sequence similarity search to first generate a set of candidates which is then rigorously refined by a series of filters exploiting the structural features specific to plant miRNAs to achieve specificity. In the first filtration step, those sequences of the candidate set whose aligned segments do not span most of the mature segment of the query are discarded. In the second filtration step, a modified Smith–Waterman pairwise alignment algorithm (Smith and Waterman 1981) is

Table 1.2 Some popular computational tools for plant miRNA identification

Tools	General features	References	Web sources
microHARVESTER	Identifies candidate miRNA homologs in any set of sequences, given a query miRNA. The candidate genes are subjected to various filters before final conclusion.	Dezulian et al. (2006)	http://ab.inf.uni-tuebingen.de/software/microHARVESTER/
MIRcheck	Uses a sequence/structure specification and 20mer coordinates. A strong conservation criteria is applied.	Jones-Rhoades and Bartel (2004)	http://web.wi.mit.edu/bartel/pub/software.html
Semrina	Uses a putative target sequence as input and allows to search for miRNAs that target this sequence.	Munoz-Merida et al. (2012)	http://www.bioinfocabd.upo.es/semirna
PsRobot	Web-based tool for identification of smRNAs with stem-loop precursors and their target genes.	Wu et al. (2012)	http://omicslab.genetics.ac.cn/psRobot/
MaturePred	A machine-learning method based on support vector machine which predicts positions of plant miRNAs for new plant pre-miRNA candidates.	Xuan et al. (2011)	http://nclab.hit.edu.cn/maturepred/
MiRDeep-P	Enables the expression patterns of annotated miRNA genes from small RNA library generated by NGS and explores the expression pattern of annotated miRNA genes.	Yang and Li (2011)	http://faculty.virginia.edu/lilab/miRDP/

(continued)

Table 1.2 (continued)

Tools	General features	References	Web sources
miRDeepFinder	A software package developed to identify and functionally analyze plant microRNAs (miRNAs) and their targets from small RNA data sets obtained from deep sequencing.	Xie et al. (2012)	http://www.leonxie.com/DeepFinder.php
miRCat	miRCat is a part of a tool kit developed for analysis of large-scale plant sRNA data sets. It identifies mature miRNAs and their precursors.	Moxon et al. (2008)	http://srna-tools.cmp.uea.ac.uk/plant/cgi-bin/srna-tools.cgi?rm=input_form&tool=mircat
C-mii	C-mii pipeline is based on combined steps and criteria from previous studies and also incorporates several tools.	Numnark et al. (2012)	http://www.biotec.or.th/isl/c-mii

applied to precisely determine the mature sequence in the candidate precursor from the optimal alignment of the query mature sequence against the corresponding segment of the BLAST hit. A candidate is discarded if the length of the mature sequences differs by >2 nt. In the third filtration step, the minimal free energy structure of the candidate sequence using RNAfold (Hofacker et al. 1994) is predicted and its putative miRNA* sequence is determined. A candidate is discarded if more than six nucleotides of its miRNA* are not predicted to form bonds with its mature miRNA (keeping in mind the 2 nt offset between miRNA and miRNA*). From a selection of all candidates that pass each filter, a multiple sequence alignment is constructed, using T-Coffee (Notredame et al. 2000), of a region that includes the miRNA, the miRNA*, and the "loop" sequence in between the miRNA and the miRNA*.

Of late, such a strategy has been extensively used for identification of miRNA genes from expressed sequence tag (EST) and a genome survey sequence (GSS) data sets of plants (Zhang et al. 2005, 2007). To further minimize the presence of false-positives, the concept of calculation of minimal free energy index (MFEI) was also introduced (Zhang et al. 2006a, b, c). The miRNA precursor sequences were reported to have higher MEFI than other coding or noncoding small RNA molecules (Zhang et al. 2006b). Using this approach, many miRNAs have been successfully identified in different plant species, such as soybean (Chen et al. 2009), maize (Zhang et al. 2006c), tomato (Yin et al. 2008), tobacco (Frazier et al. 2010), potato

(Xie et al. 2011), wheat (Han et al. 2009), mustard (Xie et al. 2007), citrus (Song et al. 2009), switch grass (Xie et al. 2010), apple (Gleave et al. 2008), cotton (Wang et al. 2012), and coffee (Akter et al. 2014).

1.3.2 MIRcheck

MIRcheck (Jones-Rhoades and Bartel 2004) uses a sequence/structure specification and the coordinates of the 20mer within the hairpin. It then uses a series of metrices concerning the number of unpaired nucleotides and bulges in the miRNA mature regions and the length of the hairpin. Sequences overlapping repetitive elements are eliminated, and a strong conservation criterion is applied, retaining only those stem-loops where the mature miRNA appears in both the genomes and exhibit high conservation in both miRNA and miRNA* sequence. Additionally, stem-loops are tested for robust folding, indicating that their secondary structures do not change substantially in the presence of flanking sequences. With this tool, Jones-Rhoades and Bartel (2004) identified 379 good plant miRNA candidates in 228 unique loci, of which 23 had their expression experimentally verified.

1.3.3 Semrina

This tool takes a putative target sequence such as a messenger RNA (mRNA) as input, and allows users to search for miRNAs that target this sequence (Munoz-Merida et al. 2012). It can also be used to determine whether small RNA sequences from massive sequencing analysis represent true miRNAs and to search for miRNAs in new genomes using homology. Semrina has shown a high level of accuracy using various test sets, and gives users the ability to search for miRNAs with several different adjustable parameters. Semrina is a user-friendly and intuitive web server. It is useful in searching for miRNAs involved in particular pathways as well as those in newly sequenced genomes.

1.3.4 PsRobot

PsRobot is a web-based easy-to-use tool dedicated to the identification of smRNAs with stem-loop-shaped precursors (such as microRNAs and short hairpin RNAs) and their target genes or transcripts (Wu et al. 2012). It performs fast analysis to identify smRNAs with stem-loop-shaped precursors among batch input data and predicts their targets using a modified Smith–Waterman algorithm. The multiple user-adjustable parameters enable the software to meet different needs of users. To facilitate better classification and functional analysis of input sequences, PsRobot integrates the expression information of input sequences in reported plant smRNA binding protein pull-down assay or mutants of major smRNA biogenesis pathway genes. For example, strong association with ARGONAUTE1 (AGO1) protein and

expression depletion in dcl1 mutant will strengthen the confidence of a smRNA with stem-loop precursor as a miRNA. It also incorporates the available mRNA degradome data for users to evaluate the reliability of miRNA target prediction results. The multiplicity of miRNA binding sites on a single target as well as the cross-species conservation status of the target sites are also analyzed and provided. PsRobot can be either used online or downloaded and installed locally. The local version offers a larger capacity for input data size and has the function to incorporate user-uploaded degradome data. PsRobot has been found to outperform many of the tools in plant miRNA and target prediction.

1.3.5 MaturePred

MaturePred, a machine-learning method based on support vector machine, has been developed to predict the positions of plant miRNAs for the new plant pre-miRNA candidates (Xuan et al. 2011). Since the plant pre-miRNAs are cleaved into the miRNA:miRNA* duplexes, the prediction model considers the position-specific features for the whole miRNA:miRNA* duplexes. A set of informative features are selected to improve the prediction accuracy. Two-stage sample selection algorithm is proposed to combat the serious imbalance problem between real and pseudo miRNA:miRNA* duplexes. Extraction of the informative features is the key for improved performance of our SVM-based prediction model. MaturePred considers not only the position-specific features of a single nucleotide but also the structure-related, energy-related, and stability-related features, totaling 160 features. Some of the selected features are discussed below-.

A miRNA* is defined to have the same size as the miRNA candidate but lies on the opposite strand with its 3′ end starting two nucleotides before the matching position of the miRNA candidate's 5′ end. In order to obtain the miRNA:miRNA* candidates, two windows slide with step 1 in a pre-miRNA. If the sequence in the sliding window 1 is regarded as a miRNA candidate, the sequence in the sliding window 2 is regarded as the corresponding miRNA* candidate. The combination of windows 1 and 2 is a miRNA:miRNA* candidate. When the starting position of the miRNA candidate is coincident with the starting position of the actual miRNA, the miRNA:miRNA* candidate is a real miRNA:miRNA* duplex. Otherwise, the candidate is a pseudo miRNA:miRNA* duplex. The other information used is the stability of 5′ end of miRNA. According to miRNA biogenesis, the 5′ end of a miRNA is usually less stable than that of the corresponding miRNA*. It is useful for determining the functional strands where the miRNAs locate. Therefore, the stability of the first nucleotide at the 5′ end of miRNA/miRNA* is considered and denoted as miRNA_5′end and miRNA*_5′end, respectively. When the first position is (A, L), (G, L), (C, L), or (U, L), the feature (miRNA_5′end/miRNA*_5′end) value is assigned to 0. When it is (G, M) or (U, M), and there is a G-U or U-G wobble pair, the feature value is assigned to 1. When it is (A, M) or (U, M), and there is an A-U or U-A pair, the feature value is assigned to 2. When it is (G, M) or (C, M), and there is a G-C or C-G pair, the feature value is assigned to 3. The real miRNA:miRNA*

duplexes typically are of greater binding stability and are less likely to be broken, and MaturePred uses this information as well during the training part.

Efficacy of MaturePred has been demonstrated by Xuan et al. (2011) while working with the real miRNA:bmiRNA* duplexes and the pseudo miRNA:miRNA* duplexes in the experimentally verified pre-miRNAs (1366) from database miRBase 14 (http://www.mirbase.org/). They demonstrated the importance of careful feature extraction, feature selection, and training sample selection in achieving effective prediction performance. Particularly, according to the characteristics of plant miRNAs, 160 features were extracted and 86 informative features were selected. Each negative sample (pseudo miRNA:miRNA* duplex) was mapped into the 86-dimensional space. About 17,803 representative negative samples were selected as the training samples to combat the class imbalance problem between the positive and negative samples. The proposed two-stage sample selection method can also be applied to other class imbalance problem in bioinformatics, such as identifying the SNP sites in the EST sequences.

1.3.6 miRDeep-P

Based on ultradeep sampling of small RNA libraries by next-generation sequencing, miRDeep-P enables users to explore expression patterns of annotated miRNA genes and discover novel ones (Yang and Li 2011). To run this application, the reads should be preprocessed by removing adapters, discarding reads shorter than 15 nucleotides and parsing them into FASTA format with their copy number recorded. With correctly formatted input files, miRDeep-P maps the reads to the reference (either genomic or transcriptomic) sequences using Bowtie (Langmead et al. 2009). For a given mapped read, the optimal size of the window from which to extract reference sequences for predicting RNA secondary structure is needed. However, miRDeep-P contains a module for users to empirically determine what window sizes to use in case a set of validated miRNA genes is available. The secondary structures of the extracted reference sequences along with all reads mapped to such sequences are processed by the miRDeep core algorithm (Friedlander et al. 2008), with a plant-specific scoring system. The output from the core algorithm is then filtered with additional plant-specific criteria based on known characteristics of plant miRNA genes (Meyers et al. 2008). The overall process quantifies the signature distribution of small RNA reads and thereby provides reliable information.

1.3.7 miRDeepFinder

miRDeepFinder is a software package developed to identify and functionally analyze plant microRNAs (miRNAs) and their targets from small RNA data sets obtained from deep sequencing (Xie et al. 2012). The functions available in miRDeepFinder include preprocessing of raw data, identifying conserved miRNAs, mining and classifying novel miRNAs, miRNA expression profiling, predicting miRNA

targets, and gene pathway and gene network analysis involving miRNAs. The fundamental design of miRDeepFinder is based on miRNA biogenesis, miRNA-mediated gene regulation and target recognition, such as perfect or near-perfect hairpin structures, different read abundances of miRNA and miRNA*, and targeting patterns of plant miRNAs. To test the accuracy and robustness of miRDeepFinder, a small RNA deep sequencing data set of *Arabidopsis thaliana* published in the GEO database of NCBI was explored (Xie et al. 2012). The test retrieved 128 of 131 (97.7%) known miRNAs that have a more than 3 read count in *Arabidopsis*. Because many known miRNAs are not associated with miRNA*s in small RNA data sets, miRDeepFinder was also designed to recover miRNA candidates without the presence of miRNA*. To mine as many miRNAs as possible, miRDeepFinder allows users to compare mature miRNAs and their miRNA*s with other small RNA data sets from the same species. Cleaveland software package was also incorporated into miRDeepFinder for miRNA target identification using degradome sequencing analysis. Using this new computational tool, Xie et al. (2012) identified 13 novel miRNA candidates with miRNA*s from *Arabidopsis* and validated 12 of them experimentally. Interestingly, of the 12 verified novel miRNAs, a miRNA named AC1 spans the exons of two genes (UTG71C4 and UGT71C3). Both the mature AC1 miRNA and its miRNA* were also found in four other small RNA data sets.

1.3.8 miRCat

miRCat is a part of tool kit developed for analysis of large-scale plant sRNA data sets (Moxon et al. 2008). It identifies mature miRNAs and their precursors. Users upload a FASTA file of sRNA sequences mapped to a plant genome using PatMaN (Prüfer et al. 2008) and grouped into loci. The program screens a number of empirical criteria for bona fide miRNA loci. Briefly, it searches for a two-peak alignment pattern of sRNAs on one strand of the locus and assesses the secondary structures of a series of putative precursor transcripts using the RNAfold and randfold. miRCat has been tested on several high-throughput plant sRNA data sets and shows a high level of sensitivity and specificity.

1.3.9 C-mii

This software tool is developed for plant miRNA and target identification (Numnark et al. 2012). C-mii pipelines are based on combined steps and criteria from previous studies and also incorporated with several tools such as standalone BLAST and UNAFold and preinstalled databases including miRBase, UniProt, and Rfam. C-mii provides the following distinguished features. First, it comes with graphical user interfaces of well-defined pipelines for both miRNA and target identifications with reliable results. Second, it provides a set of filters allowing users to reduce the number of results corresponding to the constraints in plant miRNA and target biogenesis. Third, it extends the standard computational steps of miRNA target identification

with miRNA-target folding module and GO annotation. Fourth, it supplies the bird's-eye views of the identification results with infographics and grouping information. Fifth, it provides helper functions for database update and autorecovery to ease system usage and maintenance. Finally, it supports multiproject and multithread management to improve the computational speed.

1.4 Conclusion

Biocomputational tools have enabled the discovery of novel biomolecules and their interactions in complex data sets. Cracking the complete mystery of miRNA is critical to understanding the developmental and regulatory aspects of plant systems. A number of methods, as summarized in this review, have been developed in an attempt to identify plant miRNAs. It is very difficult to accurately compare the methods because they use different data sets and may be based on different parameters to evaluate their performance. Selection of a methodology to use for a study probably depends on the information available. Further, some tools generate huge false-positive data, while the majority do not provide insights into the function or regulatory role of the predicted candidates. The lack of a clear and simpler pipeline to predict and validate miRNA candidates also makes the task of predicting miRNA transcripts and their encoded miRNAs complicated. The presence of various small RNA populations in plants aggravates the challenge. Ongoing improvements in computational tools as well as laboratory techniques are quite promising for achieving breakthrough levels in miRNA research in the near future.

References

Akter A, Islam MM, Mondal SI, Mahmud Z (2014) Computational identification of miRNA and targets from expressed sequence tags of coffee (*Coffea arabica*). Saudi Arabia J Biol Sci 21:3–12

Altschul SF, Gish W, Miller W et al (1990) Basic local alignment search tool. J Mol Biol 215(3):403–410

Ambros V (2001) microRNAs: tiny regulators with great potential. Cell 107(7):823–826

Ambros V, Bartel B, Bartel DP et al (2003) A uniform system for microRNA annotation. RNA 9:277–279

Bartel B, Bartel DP (2003) MicroRNAs: at the root of plant development? Pl Physiol 132:709–717

Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism and function. Cell 116:281–297

Bonnet E, Wuyts J, Rouze P, Van de Peer Y (2004) Detection of 91 potential conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* and identification of important target genes. Proc Natl Acad Sci U S A 101:11511–11516

Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M et al (2008) Widespread translational inhibition by plant miRNAs and siRNAs. Sci 320(5880):1185–1190

Carrington JC, Ambros V (2003) Role of microRNAs in plant and animal development. Sci 301:336–338

Chambers C, Shuai B (2009) Profiling microRNA expression in Arabidopsis pollen using microRNA array and real time PCR. BMC Pl Biol 9(87):1–11
Chen R, Hu Z, Zhang H (2009) Identification of microRNA in wild soybean (*Glycine soja*). J Integr Plant Biol 51(12):1071–1079
Cullen BR (2011) Viruses and microRNAs: RISCy interactions with serious consequences. Genes Dev 25:1881–1894
Dezulian T, Remmert M, Palatnik JF et al (2006) Identification of plant homologs. Bioinformatics 22:359–360
Ding Y, Chen Z, Zhu C (2011) Microarray-based analysis of cadmium-responsoive microRNAs in rice. J Exp Biol 62(10):1–11
Fahlgren N, Howell MD, Kasschau KD et al (2007) High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of miRNA genes. PLoS One 2:e219
Frazier TP, Xie F, Freistaedter A et al (2010) Identification and characterization of microRNAs and their target genes in tobacco (*Nicotiana tabacum*). Planta 232:1289–1308
Friedlander MR, Chen W, Adamidi C, Maaskola J et al (2008) Discovering microRNAs from deep sequencing data using miRDeep. Nat Biotechnol 26:407–415
Gleave AP, Ampomah-Dwamena C, Berthold S, Dejnoprat S et al (2008) Identification and characterisation of primary microRNAs from apple (*Malus domestica* cv. Royal Gala) expressed sequence tags. Tree Genet Genomes 4:343–358
Han Y, Luan F, Zhu H, Shao Y et al (2009) Computational identification of microRNAs and their targets in wheat (*Triticum aestivum* L.) Sci China C Life Sci 52:1091–1100
He C, Cui K, Zhang J, Duan A, Zeng Y (2013) Next-generation sequencing-based mRNA and microRNA expression profiling analysis revealed pathways involved in the rapid growth of developing culms in *Moso bamboo*. BMC Pl Biol 13(119):2–14
Hofacker IL, Fontana W, Stadler PF, Bonhoeffer LS et al (1994) Fast folding and comparison of RNA secondary structures. Monatsh Chem 125:167–188
Hu LL, Huang Y, Wang QC, Zou Q, Jiang Y (2012) Benchmark comparison of ab initio microRNA identification methods and software. Genet Mol Res 11(4):4525–4538
Huang Y, Zou Q, Wang SP, Tang SM et al (2011) The discovery approaches and detection methods of micro-RNAs. Mol Biol Rep 38:4125–4135
Jiao Y, Song W, Zhang M, Lai J (2011) Identification of novel maize miRNAs by measuring the precision of precursor processing. BMC Pl Biol 11:141. doi:10.1186/1471-2229-11-141
Jones-Rhoades MW, Bartel DP (2004) Computational identification of plant MicroRNAs and their targets, including a stress-induced miRNA. Mol Cell 14:787–799
Kim J, Jung JH, Reyes JL, Kim YS et al (2009) microRNA directed cleavage of ATHB15 mRNA regulated vascular development in Arabidopsis inflorescence stems. Pl J 42:84–94
Kurihara Y, Watanabe Y (2004) *Arabidopsis* Micro-RNA biogenesis through dicer like 1 protein functions. Proc Natl Acad Sci U S A 101:12753–12758
Lagos-Quintana M, Rauhat R, Lendeckel W, Tueshi T (2001) Identification of novel genes coding for small expressed RNAs. Sci 294:853–858
Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25
Lau NC, Lim LP, Weinstein EG, Bartel DP (2001) The abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. Sci 294:858–862
Lee RC, Feinbaum RL, Ambros V (1993) The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementary to lin-14. Cell 75:843–854
Lee Y, Jeaon K, Lee JT, Kim S, Kim VN (2002) MicroRNA maturation: stepwise processing and subcellular localization. Eur Mol Biol J 21:4663–4670
Lee Y, Kim M, Han J, Yeom KM, Lee S, Baek SH, Kim VN (2004) MicroRNA genes are transcribed by RNA polymerase II. The EMBP J 23:4051–4060
Li L, Xu J, Yang D, Tan X, Wang H (2010) Computational approaches for microRNA studies: a review. Mamm Genome 21:1–12
Liu H, Tian X, Li Y, Wu C, Zheng C (2008) Microarray-based analysis of stress-regulated microRNAs in *Arabidopsis thaliana*. RNA 14:836–843

Llave C, Kasschau KD, Rector MA, Carrington JC (2002) Endogenous and silencing-associated small RNAs in plants. The Pl Cell 14:1605–1619
Lu S, Sun YH, Shi R, Clark C et al (2005) Novel and mechanical stress-responsive microRNAs in *Pouplas trichocarpa* that are absent from *Arabidopsis*. Plant Cell 17:2186–2203
Mendes ND, Freitas AT, Sagot MF (2009) Current tools for the identification of miRNA genes and their targets. Nucleic Acids Res 37:2419–2433
Meyers BC, Axtell MJ, Bartel B, Bartel DP et al (2008) Criteria for annotation of plant microRNAs. The Pl Cell 20:3186–3190
Millar AA, Waterhouse PM (2005) Plant and animal microRNAs: similarities and differences. Funct Integ Genomics 5(3):129–135
Moxon S, Schwach F, Dalmay T, MacLean D et al (2008) A toolkit for analyzing large scale plant small RNA datasets. Bioinformatics 24(19):2252–2253
Munoz-Merida A, Perkins JR, Viguera E et al (2012) Semirna: searching for plant miRNAs using target sequences. OMICS 16(4):168–177
Notredame C, Higgins DG, Heringa J (2000) T-coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol 302:205–217
Numnark S, Mhuantong W, Ingsris S, Wang X, Wichadakul D (2012) C-mii: a tool for plant miRNA and target identification. BMC Genomics 13(7):S16
Park W, Li J, Song R, Messing J, Chen X (2002) CARPEL FACTORYU, a dicer homolog and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. Curr Biol 12:1484–1495
Pasaquinelli AE, Reinhart BJ, Slack F et al (2000) Conservation of the sequence and temporal expression of let-7 heterochromic regulatory RNA. Nature 408:86–89
Prüfer K, Stenzel U, Dannemann M, Green RE et al (2008) PatMaN: rapid alignment of short sequences to large databases. Bioinformatics 24:1530–1531
Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE et al (2000) The 21-nucleotide let-7 RNA regulates developmental timing in *Caernorhabditis elegans*. Nature 403:901–906
Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP (2002) MicroRNAs in plants. Genes Dev 16:1616–1626
Smith TF, Waterman MS (1981) Identification of common molecular subsequences. J Mol Biol 147:195–197
Song C, Fang J, Li X, Liu H, Thomas C (2009) Identification and characterization of 27 conserved microRNAs in citrus. Planta 230:671–685
Sunkar R, Girke T, Jain PK, Zhu J (2005) Cloning and characterization of microRNAs from Rice. Pl Cell 17(5):1397–1411
Unver T, Namuth-Covert DM, Budak H (2009) Review of current methodological approaches for characterizing microRNAs in plants. Int J Pl Genomics. doi:10.1155/2009/262463
Wang M, Wang Q, Wang B (2012) Identification and characterization of MicroRNAs in Asiatic cotton (*Gossypium arboreum* L.) PLoS One 7(4):e33696
Wu H, Ma Y, Chen T, Wang M, Wang X (2012) PsRobot: a web-based plant small RNA meta-analysis toolbox. Nucleic Acids Res 40:W22–W28. doi:10.1093/nar/gks554
Xie F, Frazier TP, Zhang B (2010) Identification and characterization of microRNAs and their targets in the bioenergy plant switchgrass (*Panicum virgatum*). Planta 232:417–434
Xie F, Frazier TP, Zhang B (2011) Identification, characterization and expression analysis of MicroRNAs and their targets in the potato (*Solanum tuberosum*). Gene 473:8–22
Xie F, Huang SQ, Guo K, Xiang AL et al (2007) Computational identification of novel microRNAs and targets in *Brassica napus*. FEBS Lett 581:1464–1474
Xie F, Xiao P, Chen D, Xu L, Zhang B (2012) miRDeepFinder: a miRNA analysis tool for deep sequencing of plant small RNAs. Pl Mol Biol 80(1):75–84
Xie Z, Allen E, Fahlgren N, Calamar A, Givan SA, Carrington JC (2005) Expression of Arabidopsis MIRNA genes. Pl Physiol 138(4):2145–2154
Xuan P, Guo M, Huang Y, Li W, Huang Y (2011) MaturePred: efficient identification of MicroRNAs within novel plant pre-MiRNAs. PLoS One 6(11):e27422. doi:10.1371/journal.pone.0027422

Yang X, Li L (2011) miRDeep-P: a computational tool for analyzing the microRNA Transcriptome in plants. Bioinformatics 1:1–2

Yin Z, Li C, Han X, Shen F (2008) Identification of conserved microRNAs and their target genes in tomato (*Lycopersicon esculentum*). Gene 414:60–66

Yu B, Yang Z, Li J, Minakhina S, Yang M et al (2005) Methylation as a crucial step in plant microRNA biogenesis. Sci 307(5711):932

Zhang BH, Pan XP, Anderson TA (2006a) Identification of 188 conserved maize microRNAs and their targets. FEBS Lett 580:3753–3762

Zhang BH, Pan XP, Cox SB, Cobb GP, Anderson TA (2006b) Evidence that miRNAs are different from other RNAs. Cell Mol Life Sci 63:246–254

Zhang BH, Pan XP, Wang QL et al (2005) Identification and characterization of new plant microRNAs using EST analysis. Cell Res 15(5):336–360

Zhang BH, Pan X, Cobb GP, Anderson TA (2006c) Plant microRNA: a small regulatory molecule with big impact. Dev Biol 289:3–16

Zhang BH, Wang Q, Wang K, Pan X et al (2007) Identification of cotton microRNAs and their targets. Gene 397:26–37

Zhang Z, Yu J, Li D, Zhang Z et al (2010) PMRD: plant microRNA database. Nucleic Acids Res 38(Database issue):D806–D813

Zhao M, Tai H, Sun S, Zhang F et al (2012) Cloning and characterization of maize miRNAs involved in response to nitrogen deficiency. PLoS One 7(1):e29669

Zhou JS, Guan J (2012) Finding MicroRNA targets in plants: current status and perspectives. Genomics Proteomics Bioinformatics 10:264–275

Control of Gene Expression by RNAi: A Revolution in Functional Genomics

2

Suresh Kumar and Raj Kumar Salar

Abstract

RNA interference is an evolutionarily preserved gene-silencing phenomenon in eukaryotic organisms, which is triggered by double-stranded RNA (dsRNA) derived from viral replication intermediates or transposable elements or an introduced synthetic oligonucleotide. Dicer, an endonuclease, processes dsRNA into 21- to 23-base-pair (bp) small interfering RNA (siRNA) with 2-nucleotide 3′ overhangs, which are further linked with the RNA-induced silencing complex (RISC). Argonaute, an active (AGO)-family protein of RISC remove the sense strand of siRNA, guided by the antisense strand, mediates the destruction of cognate target RNA with exquisite potency and selectivity. RNAi is employed to downregulate the activity of a gene with hereditability, and other significant advantages have revolutionized functional genomics to discover genes involved in disease progression and validate gene function, holding promise for the expansion of new therapeutics to treat several diseases. RNAi has enormous exciting applications in plant biotechnology, generating new hopes to produce flowers with desired colors, and to enhance crop yield and shelf life, while in humans it is used to treat various forms of cancer, infectious diseases, and neurodegenerative disorders, though the stability, off-target activity, and in vivo delivery of small interfering RNAs remain significant obstacles. This chapter reviews the mechanism of RNA interference, its applications, and challenges encountered while using this technology for research purposes.

S. Kumar • R.K. Salar (✉)
Department of Biotechnology, Chaudhary Devi Lal University, Sirsa 125 055, Haryana, India
e-mail: rajsalar@rediffmail.com

S.K. Gahlawat et al. (eds.), *Plant Biotechnology: Recent Advancements and Developments*, DOI 10.1007/978-981-10-4732-9_2

2.1 Introduction

The central dogma of molecular biology describes the process by which the genetic information in DNA is converted into a functional product. In eukaryotic cells, there are several steps in the pathway leading from DNA to protein, and all of them can be controlled (Fig. 2.1). Eukaryotic cells can control the expression of a gene through (1) control of when and how often a given gene is transcribed; (2) processing of primary RNA transcripts; (3) transport of selected mRNAs from the nucleoplasm to the cytoplasm; (4) systemic degradation of selective mRNAs; (5) selection of which mRNAs are translated by ribosomes; and (6) selective activation or inactivation of posttranslational proteins. Gene expression can be regulated at six mentioned steps; however, for most genes, the control of transcription, posttranscriptional processing, translation, and posttranslational modification is paramount.

The regulation of transcription and translation has long been considered the "Holy Grail" of molecular biology (Smith 2009). Researchers have made exciting progress in developing a novel and effective approach to regulate transcription and translation activity of any gene. They have accomplished this by introducing a synthetic oligonucleotide molecule, which triggers a cellular process called *RNA interference* (RNAi). The basic goal of RNAi or antisense is to find or design small complementary oligonucleotides and get them inside a cell where they can bind to a segment of mRNA, prevent translation, or initiate cleavage of mRNA to interrupt protein synthesis. Probably, RNAi is the earliest immune protecting mechanism that enables organisms to shield their genome from foreign nucleic acids such as transposons and viruses (Waterhouse et al. 2001). It also refers to the accurate influence of DNA or RNA so that they transcribe to form a complimentary copy of normal—or *sense*—messenger RNA (mRNA). The antisense strand will target and bind with a specific mRNA of a desired gene to inhibit its phenotypic expression via different mechanisms (Fig. 2.2). This inhibition is sometime called *knockdown* or *knockout* subject to the gene function being either partially or completely abolished; it also permits researchers to set the usefulness of the designed sequence.

RNAi was initially reported in transgenic plants as cosuppression, a mechanism of gene suppression similar to that of transcriptional gene silencing (TGS) or

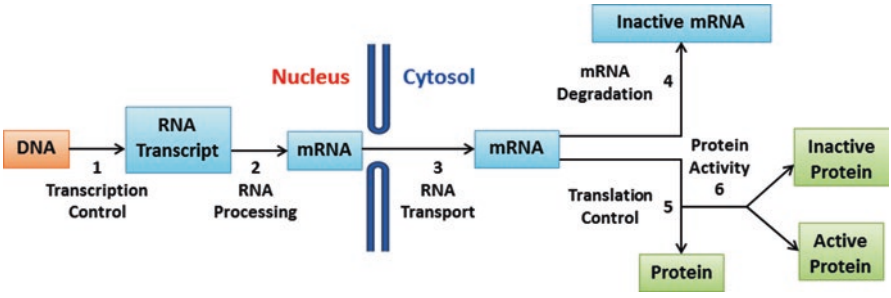

Fig. 2.1 Eukaryotic gene expression can be controlled at six different steps (Adapted from Bruce et al. 2013)

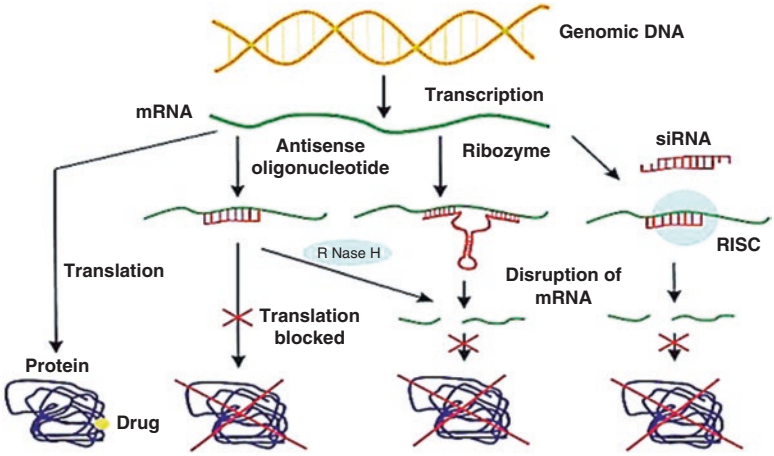

Fig. 2.2 Different gene expression–controlling techniques. Most conventional drugs bind to particular sites of a protein to inhibit their expression, whereas antisense strands bind with their complementary target mRNA and block translation of that mRNA (antisense oligonucleotides) or may induce mRNA degradation by RNase H/ribozymes. The RNAi approach is accomplished with siRNA/miRNA having 2-nucleotide 3′ overhangs that are associated with RISC and trigger degradation of the target mRNA (Adapted from Kurreck 2003)

posttranscriptional gene silencing (PTGS). This technique refers to several distinct but related approaches such as classical antisense, ribozyme, quelling (Lindbo 2012), and aptamer technology (Zhang and Wang 2005) (Table 2.1).

RNAi is a posttranscriptional method triggered with the introduction of double-stranded RNA (dsRNA) either through a virus or transgenes homologous to the target gene facilitating sequence-specific gene silencing. Dicer, a ribonuclease III enzyme processes long double-stranded RNAs to generate duplexes of ≈21 nt with 2-nucleotide 3′- overhangs called *small interfering RNA* (siRNA) that induce sequence-specific messenger RNA (mRNA) degradation. RNAi has been reported to naturally occur in organisms as diverse as bacteria, fungi, algae, plants, and animals (invertebrates and vertebrates) (Table 2.2).

The first evidence of gene silencing via dsRNA was reported successfully in roundworm, *Caenorhabditis elegans*, and this phenomenon has been termed *RNA interference* (RNAi) (Fire et al. 1998). In 2006, Fire and Mello shared the Nobel Prize in Physiology and Medicine for their innovative work on RNAi. Further studies in fruit fly, *Drosophila melanogaster*, have contributed greatly toward comprehension of the biochemical nature of the RNAi pathway (Elbashir et al. 2001). RNAi has become an invaluable research tool in both in vitro and in vivo studies because synthetic dsRNA introduced into cells can selectively and robustly induce suppression of specific genes of interest. This chapter reviews the RNAi mechanism, its applications, and challenges with the use of this technology for research purposes.

Table 2.1 Characteristics of gene expression–controlling technologies (After Zhang and Wang 2005)

Technology	Active molecule	Molecular target	Cellular site of action[a]	Suggested mechanism of action	Status of drug discovery and development
Antisense	DNA or RNA	RNA[b]	Cytoplasm	Translation blockage, RNase H activation, inhibition of splicing, disruption of RNA structure	Clinical use[c], clinical trials
RNAi	RNA	RNA	Cytoplasm	Promotion of mRNA degradation	In vivo and in vitro study
Ribozyme	RNA	RNA	Cytoplasm	Translation inhibition, destruction of RNA structure	Clinical use
Triplex forming oligonucleotide	DNA	DNA	Nucleus	Transcription inhibition	Preclinical
Aptamer	DNA or RNA	Protein	Nucleus, cytoplasm	Interference with protein function	Clinical phase I trial
Riboswitch	RNA	RNA	Nucleus, cytoplasm	Inhibition or promotion of gene expression at the level of transcription, translation, or RNA splicing	Clinical use, clinical trials
Quelling	RNA	RNA	Nucleus, cytoplasm	Induction of silencing of transgenes and cognate endogenous mRNAs at the posttranscriptional level	Clinical use, clinical trials

[a]Site of action refers to the initial site of interaction of the active molecule with its target molecule and may not be the same as the site where biological activity occurs
[b]Any kind of RNA can be targeted: pre-RNA, mRNA, or viral RNA
[c]The first antisense drug, Vitravene, has been approved for the treatment of patients with cytomegalovirus-induced retinitis

2.2 Historical Perspectives

RNAi was initially discovered in transgenic plants where antisense RNAs prevent transcription of a peculiar gene and indirectly from observing unpredicted results in experiments performed to obtain transgenic petunias with a higher content of anthocyanin pigments by introducing additional copies of chalcone synthase transgene (Napoli et al. 1990; Jorgensen et al. 1996). To the researchers' surprise, they obtain variegated or chimeric flowers instead of intense purple coloration (Fig. 2.3).