







JOURNAL OF ZHEJIANG UNIVERSITY

AGRICULTURE AND LIFE SCIENCES

浙江大学学报

农业与生命科学版

ISSN 1008-9209



4/2014

Vol.40 No.4 July 2014

中国·杭州 HANGZHOU, CHINA Zhejiang Daxue Xuebao (Nongye yu Shengming Kexue Ban)

《浙江大学学报(农业与生命科学版)》编委会

Editorial Board of Journal of Zhejiang University (Agriculture & Life Sciences)

主编 Editor-in-Chief

军(Zhu Jun, Zhejiang University, jzhu@zju.edu.cn)

副主编 Associate Editor-in-Chief

吴 平(Wu Ping), Zhejiang University)

国外编委 Foreign Members of the Editorial Board

Chrisman Cheryl L., University of Florida, USA

Kohanawa Masashi, Hokkaido University, Japan

Lai Chaoqiang, Tufts University, USA

May Sean, University of Nottingham, UK

Qin Jianguang, Flinders University, Australia

Song Guoqing, Michigan State University, USA

Stafford James L., University of Alberta, Canada

Tang Lie, Iowa State University, USA

Whelan James M., University of Western Australia, Australia

Xie Jiahua, North Carolina Central University, USA

Zhao Fengqi, University of Vermont, USA

Zhu Qianhao, CSIRO Plant Industry, Australia

国内编委 Domestic Members of the Editorial Board

包劲松(Bao Jinsong, Zhejiang University)

ヲ宏燕(Diao Hongyan, Zhejiang University)

方维焕(Fang Weihuan, Zhejiang University)

管敏鑫(Guan Minxin, Zhejiang University)

勇(He Yong, Zhejiang University)

胡元亮(Hu Yuanliang, Nanjing Agricultural University)李 铎(Li Duo, Zhejiang University)

李民赞(Li Minzan, China Agricultural University)

刘建新(Liu Jianxin, Zhejiang University)

刘向东(Liu Xiangdong, Zhejiang Sci-Tech University) 马忠华(Ma Zhonghua, Zhejiang University)

缪云根(Miao Yungen, Zhejiang University)

汪俏梅(Wang Qiaomei, Zhejiang University)

吴信忠(Wu Xinzhong, Zhejiang University)

叶兴乾(Ye Xingqian, Zhejiang University)

章明奎(Zhang Mingkui, Zhejiang University)

赵宇华(Zhao Yuhua, Zhejiang University)

陈 铭(Chen Ming, Zhejiang University)

樊龙江(Fan Longjiang, Zhejiang University)

傅 衍(Fu Yan, Zhejiang University)

郭仰东(Guo Yangdong, China Agricultural University)

胡松华(Hu Songhua, Zhejiang University)

梁月荣(Liang Yuerong, Zhejiang University)

刘维屏(Liu Weiping, Zhejiang University)

王 俊(Wang Jun, Zhejiang University)

岩(Wang Yan, Zhejiang University) 杨肖娥(Yang Xiaoe, Zhejiang University)

张国平(Zhang Guoping, Zhejiang University)

张子丁(Zhang Ziding, China Agricultural University)

郑 平(Zheng Ping, Zhejiang University)

周东美(Zhou Dongmei, Chinese Academy of Sciences) 朱伟云(Zhu Weiyun, Nanjing Agricultural University)

执行编辑 Managing Editor

胡清华(Hu Qinghua, zdxbnsb@zju.edu.cn)

浙江大学学报

(农业与生命科学版)

Zhejiang Daxue Xuebao

(Nongye yu Shengming Kexue Ban)

第 40 卷 第 4 期(总第 185 期) 2014 年 7 月 20 日

ISSN 1008-9209 CN 33-1247/S JOURNAL OF ZHEJIANG UNIVERSITY

(Agriculture & Life Sciences)

Vol. 40 No. 4 (General serial No. 185) July 20, 2014 http://www.journals.zju.edu.cn/agr

目 次 Contents

本刊编委会委员名单
稿 约
作物基因组与数量遗传学专辑(Special issue on crop genome and quantitative genetics)
综述(Reviews)
高通量测序技术在多倍体作物基因组学研究中的应用(英文) 朱乾浩,丹尼.卢埃琳,印.威尔逊(355
Use of next-generation sequencing in genomic studies of polyploid crops: cotton as an example
植物非编码小 RNA 研究进展(英文) ····································
Recent studies on non-coding small RNAs in plants Shen Enhui, Liu Yang, Ye Chuyu, Fan Longjiang (370
数量性状位点定位和全基因组关联分析的方法与软件综述(英文) 马蒙,朱军(379
Tools for quantitative trait locus mapping and genome-wide association study mapping: a review
研究论文(Research articles)
基于四交群体的连锁图谱构建(英文) 蒋蓓蓓,余世洲,肖炳光,楼向阳,徐海明(387
Constructing linkage map based on a four-way cross population
Jiang Beibei, Yu Shizhou, Xiao Bingguang, Lou Xiangyang, Xu Haiming (387
东乡野生稻叶绿体基因组拼接及系统进化分析(英文) 林张翔,王营营,付菲,叶楚玉,樊龙江(397
Complete chloroplast genome of Dongxiang wild rice and its application in phylogenetic analysis
Lin Zhangxiang, Wang Yingying, Fu Fei, Ye Chuyu, Fan Longjiang (397
茶树"龙井 43"叶绿体基因组测序及其系统进化(英文)
Entire chloroplast genome sequence of tea (Camellia sinensis cv. Longjing 43): a molecular phylogenetic analysis
······ Ye Xiaoqian, Zhao Zhonghui, Zhu Quanwu, Wang Yingying, Lin Zhangxiang, Ye Chuyu, Fan Longjiang, Xu Hairong (404
吸烟对控制 FEV1/FVC 比率的基因上位性及性别互作效应的影响(英文) 许昌巍,朱军(413
Impacts of cigarette smoking on epistasis and gender-specific effects of FEV1/FVC ratio in human
吸烟和性别对身体质量指数遗传结构的影响(英文)
Impact of cigarette smoking and gender on genetic architecture of body mass index
今其田纽关联分析 SCDO 港传亦見与昭悃的关系(黄文)

Genome-wide association studies for identifying genetic architecture of SGRQ in smoking population
Hao Xinying, Zhu Jun (431)
烟草核心种质库构建及遗传多样性研究(英文)
Genetic diversity and development of core collection in tobacco (Nicotiana tabacum L.) resources
······ Zhou Jiaping,
Yang Chunyuan, Wu Chun, Wang Rengang, Shi Yuewei, Xie Shengdong, Wang Zhihong, Xu Haiming, Ren Xueliang (440)
玉米 C 型细胞质雄性不育系及保持系基因差异表达分析 ····································
Different expression analysis of genes between C type cytoplasmic male sterile line and its maintainer
line in maize ······ Wang Jiyue, Wang Shengqing, Cao Moju (451)
水稻 Os09g24220 基因插入突变体的分子鉴定与农艺性状分析
Molecular characterization and agronomic trait analysis of rice Os09g24220 gene insertion mutants
建兰转录本的微卫星序列和单核苷酸多态性信息分析 李小白,向林,罗洁,秦德辉,孙崇波(463)
Analysis of microsatellite and single nucleotide polymorphism within transcriptomic database in
Cymbidium ensifolium ····· Li Xiaobai, Xiang Lin, Luo Jie, Qin Dehui, Sun Chongbo (463)
补 白(Filler)
欢迎订阅《浙江大学学报(农业与生命科学版)》
新书架

CN33-1247/S * 1956 * b * A4 * 120 * zh+en * P * \S 20.00 * 1000 * 13 * 2014-07

责任编辑 胡清华 英文校对 I. H. Shamsi Journal of Zhejiang University (Agric. & Life Sci.)

http://www.journals.zju.edu.cn/agr

E-mail: zdxbnsb@zju. edu. cn

DOI:10.3785/j.issn.1008-9209.2014.04.211

Recent studies on non-coding small RNAs in plants

Shen Enhui, Liu Yang, Ye Chuyu, Fan Longjiang* (Zhejiang Key Laboratory of Crop Germplasm/Department of Agronomy, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310058, China)

Summary There are many non-coding small RNAs in plants and animals, which regulate gene expression via direct cleavage of target mRNAs or via inhibition of translation at the post-transcriptional level. In this paper, recent studies on microRNA (miRNA)-mediated phased siRNAs (phasiRNA) and endogenous target mimics (eTMs) were reviewed: 1) phasiRNAs can be generated both in coding and non-coding loci, and some require miRNA-mediated cleavage for their biogenesis. Pattern and evolutionary mechanisms of miRNA-mediated phasiRNAs were discussed. 2) Genome-wide identification and application of eTMs as miRNA decoy targets were recommended, including the use of artificial target mimics to validate the functions of miRNAs and employing bioinformatics methods to identify eTMs in the whole-genome level.

Key words microRNA; small interfering RNA; phased small interfering RNA; endogenous target mimics; bioinformatics

CLC number Q 943 Document code A

植物非编码小 RNA 研究进展(英文). 浙江大学学报(农业与生命科学版), 2014, 40(4): 370-378

沈恩惠, 刘扬, 叶楚玉, 樊龙江"(浙江大学农业与生物技术学院农学系/浙江省作物种质资源重点实验室, 杭州 310058)

摘要 在植物和动物中存在着大量非编码小 RNA,它们通过对靶标 mRNA 直接切割或在转录后抑制其翻译对基因表达起调控作用.本文主要综述了该领域在以下 2 方面的研究进展. 1)微 RNA(miRNA)介导的具有相位排列的小干扰 RNA(phasiRNA):由 miRNA 介导的 phasiRNA 可由编码和非编码位点产生;介绍了 phasiRNA 产生的模式特征,以及 phasiRNA 和 miRNA 的进化机制. 2)介绍了最近出现的内源 miRNA 诱捕靶标(endogenous target mimics, eTMs)研究进展,包括利用人工诱捕靶标检验 miRNA 的功能,通过生物信息学方法在全基因组水平计算识别 eTMs 等.

关键词 微 RNA; 小干扰 RNA; 相位排列的小干扰 RNA; 内源诱捕靶标; 生物信息学

There are many small non-coding RNAs in plants, which regulate gene expression through direct

cleavage of the target mRNAs or through suppression of translation at the post-transcriptional level. Small

Foundation item: Project supported by the National Natural Science Foundation of China (No. 31200117), Ministry of Science and Technology, China (No. 2013IM030700) and the Yunnan Provincial Tobacco Company Project (Nos. 08A05 and 2011YN04).

Biography: Shen Enhui, E-mail: ttxsenhui@gmail.com

Received: 2014 - 04 - 21; **Accepted:** 2014 - 05 - 19; **Published online:** 2014 - 07 - 24 URL: http://www.cnki.net/kcms/doi/10.3785/j.issn.1008-9209.2014.04.211.html

^{*} Corresponding author: Fan Longjiang, Tel: +86 -571 -88982730; E-mail: fanlj@zju.edu.cn

RNAs in plants are divided into two main categories: miRNA (microRNA) and siRNA (small interfering RNA). In this review, we present recent progress on the miRNA-mediated phased siRNAs (phasiRNA), which can be generated by cis and trans-regulation through miRNA. phasiRNA originates from transregulation, which targets more genes, and thus has received much attention recently. phasiRNA can be generated in both coding and non-coding loci with a specific pattern feature that makes it possible to develop bioinformatics methods for large-scale identification. Some studies focused the evolutionary mechanism between the miRNA and phasiRNA, which was helpful in studying the functions of phasiRNA. In addition, this review introduces the research on the recently identified endogenous target mimics (eTMs). It is known that miRNA can target genes and play an inhibitory effect on specific mRNA, however, during the plants' development, a class of long non-coding RNAs named eTMs bind to miRNA and block its functions by segregating miRNA. Currently, we are able to identify the eTMs at the whole-genome level.

1 Researches about miRNA-mediated phasiRNA

1.1 phasiRNA

In plants, some target sites or genes of miRNAs (such as miR173, miR390, miR393) can trigger the production of phased secondary siRNAs from either non-coding (such as *TAS* and miR390) or protein-coding genes (such as *AFB* and miR393). With some obvious structural features of phasiRNAs and known reference genomes, we can use bioinformatics methods to detect such the types of siRNAs with the combination of large-scale small RNA data. Howell *et al.* [1] identified the *TAS* candidate gene loci based on a remarkable feature of small RNAs that consolidates and starts the coordination of a given 21-nucleotide phase within an eight-cycle window that generates phasiRNA loci. *P* value was used as a parameter to describe the phased score, and it was influenced by

both small RNA abundance and the number of positions occupied. P values were applied in a scrolling window format with single nucleotide advances, and scores were assigned to the fourth cycle position in phasing plots. Chen et al. [2] also developed a method to detect phasiRNA, and similar to Howell et al. [1], they took into main consideration the distribution of phasiRNAs and constructed the P value to search the candidate loci with the use of a hypergeometric distribution [3]. A larger P value score indicated a more significant phased structure. In addition, Rajagopalan et al. [4] have the same core principle but a looser standard that detects the phased loci with the permission of a two base difference. Based on the above methods, Liu et al. [5] identified hundreds of phasiRNA loci in wild rice, and further research found that these sites could be important under the selection pressure during rice domestication. Many phasiRNA sites have been identified in other species^[1-2,5-6], which indicates that it may be a conservative pathway to regulate the growth of plants.

miRNA-mediated siRNAs have two regulatory pathways: one is *cis*-regulatory and the other is *trans*-regulatory (Fig. 1). We pay more attention to the *trans*-regulatory siRNAs, which can target one or more downstream genes, especially the phasiRNA loci originated from protein-coding genes.

phasiRNA

Demonstrated pattern of phasing, but for which function in *cis* or *trans* is unknown; derived from a *PHAS* gene.





phasiRNAs for wich function in trans is experimentally demonstrated; derived from a TAS gene.

tasiRNA



phasiRNAs for which function in cis is experimentally demonstrated; derived from a PHAS or TAS gene?

phasiRNA-generating loci are called PHAS genes, as tasiRNA-generating loci are TAS genes[6].

Fig. 1 Definition of phased small interfering RNA classes in plants

1.2 phasiRNA can be generated from both non-coding and coding regions

1.2.1 miRNA-mediated phasiRNA from non-codingloci There are many types of siRNAs, some of

which are specific in plants with a phased array structure. Phased arrays are a phenomenon where the primary transcript produces a continuous specific length of small RNA products with the participation of some transcription factors. One typical type was tasiRNA, which is a type of siRNAs found in some research to take part in the changing development stage of leaves in the model plant *Arabidopsis*^[7-9] (Table 1). The synthesis of tasiRNA starts from a single-stranded RNA that is then cleaved by a specific miRNA. The cleavage products are formed into a double-stranded RNA by RDR6 (RNA-dependent RNA polymerase 6), and then the DCL (Dicer-like) enzymes act on complexes to catalyze the formation of small RNAs. Although the role of the binding sites with miRNA is still not clear [10], different tasiRNAs have different miRNA binding sites, such as miR173 (TAS1 and TAS2), (TAS3),miR390 and miR828 $(TAS4)^{[4,7-8,10-12]}$. Among them TAS3 is the most conservative and has been found in many terrestrial plants (such as rice)^[10,13-14]. Furthermore, the

expression of different phased siRNAs changed significantly, and parts of them can be formed into complexes by the bind of the ARGONAUTE protein to cleave the target mRNAs^[13]. Zhao et al. ^[15] found some phased arrays of siRNAs in Chlamydomonas reinhardtii but did not find RDR homologous genes in its genome; therefore, the mechanism that produced tasiRNAs was unclear. Heisel et al. [16] and Zhu et al. [17] have found a specific type of siRNAs, which is called long miRNA-like hairpin and is related to the development of rice seeds in the 6th and the 12th chromosomes. The single-stranded RNA can be formed into a long miRNA-like precursor structure, but when the loop is quite large (>1 kb), the long stem can form into 21-nucleotide phased arrays. This formation is similar to tasiRNA duplexes, and also supports the mentioned regulation of the DCL4 in rice, which can influence the formation of siRNAs. These long miRNA-like sites were also found in other plants[4,18], and they were considered to be derived from the duplication of progenitor genes^[16].

Table 1 phasiRNA identified in non-coding loci in plants

miRNA	Length of miRNA/nt	phasiRNA loci	Species	References
miR173	22	TAS1a-c, TAS2	Arabidos psis thaliana	Vazquez et al. [7]; Peragine et al. [8]; Axell et al. [10]
miR390	21	TAS3	A. thaliana	Fahlgren et al. [11]; Willams et al. [12]
miR828	22	TAS4	A. thaliana	Rajagopalan et al. [4]
miR529, miR156	21,20	TAS6	Physcomitrella patens	Arif et al. [19]

1.2.2 miRNA-mediated phasiRNA from proteincoding loci Some miRNA-mediated phasiRNAs are from protein-coding genes. Howell et al. [1] and Chen et al. [2] first reported this phenomenon (Table 2) in several resistance genes (such as NBS-LRR genes). Due to the distribution of small RNAs in the genome as hotspots or clusters, Johnson et al. [20] generated a large data set of small RNAs from different tissues and developmental stages in rice and identified a large number of phased 21-mer clusters, which were flanked by a degenerate 22-nt motif that was offset by 12-nt from the main phase of the cluster and were preferentially expressed in inflorescences with a specific pattern. The miR2118 was one member of the 22-nt motif pattern, and the 24-mer clusters also had a specific

expression pattern. In dicot, miR2118 was also discovered and was shown to target NBS-LRR disease resistance genes, and it made this gene locus generate phasiRNA, which can act as both cis and trans regulators [6,21-23]. Xia et al. [23] found that miR828 targeted MYB genes in apples, while some of the miR828-targeted MYBs can generate over 100 sequencedistinct siRNAs that potentially target over 70 diverse genes. Källman et al. [24] found that many sRNA (small RNA) sequences in spruce can be assigned to a 21-nt siRNA sequence, which originated from the degradation of NBS-LRR genes. Further results showed that this phased degradation of resistance-related genes that contained either a TIR or NBS domain was initiated from 22-nt miRNA-guided an abundant cleavage.

Intriguingly, the known miRNA-mediated phasiRNA loci, which were generated from protein-coding genes, were associated more or less with NBS-LRR resistance

genes (Table 2). It is suggested that the regulation mechanism of miRNA-mediated phasiRNA may play an important role in the regulation of plant resistance.

Table 2 phasiRNA identified in protein-coding loci in plants

miRNA	Length of miRNA/nt	t phasiRNA loci	Species	References
miR828, miR159, miR858	22,21,22	MYB	Malus domestica	Xia et al. [23]
miR161/miR400	21	PPP clade	A. thaliana	Howell et al. [1]; Chen et al. [2]
miR393	22	siRAAR	A. thaliana	Howell et al. [1]; Windels et al. [25]
miR780/miR856	21	ATCHX18	A. thaliana	Howell et al. [1]
miR472	22	NBS-LRR	A. thaliana	Howell et al. [1]
miR2109	22	NBS-LRR	Medicago truncatula	Zhai et al. [6]
miR1507	22	NBS-LRR; DCL2	M. truncatula	Zhai et al. [6]
miR1509	22	Predicted transcription factor	M. truncatula	Zhai et al. [6]
miR5754	22	Predicted protein kinase	M. truncatula	Zhai et al. [6]
miR156/miR172	21	AP2-like	M. truncatula	Zhai et al. [6]
miR4376	22	ACA10	Solanum lycopersium	Wang et al. $^{[26]}$
miR482	22	NBS-LRR	S. lycopersium	Shivaprasad et al. [21]
mi R 2118	22	NBS-LRR; SGS3	Plant	Shivaprasad et al. $^{[21]}$; Zhai et al. $^{[6]}$; Johnson et al. $^{[20]}$
miR6019	22	NBS-LRR	Nicotiana tabacum	Li et al. [27]
miR6020	21	NBS-LRR	N. tabacum	Li et al. [27]
miR7122	22	PPR	Plant	Xia et al. [28]

1.3 Pattern of miRNA-mediated phasiRNA

miRNA complementary sites exist upstream and downstream of the region that generates small RNAs, which can be divided into two types: One is a single-hit model (i. e., TAS1/2/4)^[4,7-8,10] and the other is two-hit model (i. e., TAS3/ 6) [11-12,19]. The length of miRNA in the single-hit mode was 22-nt, and 21-nt in the two-hit model; the formation direction of the single-hit mode phasiRNA started at the 5' cleavage site, and at the 3' cleavage site in the two-hit model. There were also some variants in the single-hit mode, generating 24-nt such phasiRNA (miRNA2275)[29]; similarly, some variants had miRNA complementary sites both upstream and downstream, but the 3' site was not cleavable in the two-hit model (miR1509, miR7122)^[6]. Another existing variant (TAS6) was that both 5' and 3' ends were cleavable in order to generate phasiRNA in the two-hit model. In addition, the cleavage sites of miRNA-mediated phasiRNA tended to occur in the 10th – 11th bases of the 5' end of miRNA, which made it possible to identify phasiRNA sites on a large scale.

1.4 Evolution mechanism of miRNA and phasiRNA

Learning the evolution mechanism phasiRNA and its mediated miRNA would contribute to the identification and analysis of phasiRNA. According to the sequence similarity of the binding sits between miR161/miR163 genes and their target genes in Arabidopsis, Allen et al. [18] proposed that genes encoding miRNAs in plants originated in the inverted duplication of target gene sequences. Due to progenitor sequences of miRNA, the inverted duplication events occur and result in head-to-head or tail-totail orientations of complete or partial gene sequences. With the constraints of the foldback

transcript and DCL enzymes, the miRNA genes were formed to regulate the expression of their founder genes or other members. Rajagopalan et al. [4] also found the similar sequence of six miRNA genes in Arabidopsis, excluding miR161 and miR163, between stem structure and corresponding region of the target genes. Heisel et al. [16] showed that miRNA-like hairpin genes also share a similar evolutionary mechanism and regarded the POT gene as the progenitor of Os06g21900 gene. However, the sites of tasiRNA did not have the same foldback transcript as miRNA did, and the formation of duplexes was depended on RDR rather than the DCL, indicating that the origin of tasiRNA could be different [18]. In fact, when they first discovered tasiRNA in Arabidopsis, Vazquez et al. [7] pointed out that the lack of similar sequences in the tasiRNA site and its target gene would make the tasiRNA originate from another mechanism. Axtell et al. [10] proposed a "two-hit trigger" mechanism in the study of the universally conserved TAS3 of tasiRNA genes in higher plants in which both sites were cleaved to generate two transcripts containing miR390 complementary sites, which resulted in the formation of phased siRNA. Howell et al. [1] thought that TAS genes may have originated from the regulatory gene of miRNA. PPR-P genes, as a rapidly expansive gene family, were similar to the known TAS loci. They generated many 21-nt phased small RNAs, which were responsible for the regulation of PPR-P at the posttranscriptional level. The genetic drifts or recombination events resulting in the loss of coding potential and the reservation of the small RNA binding sites and phasiRNA producing regions may have arisen as an evolutionary mechanism for the new TAS sites^[1]. Shen et al. [30] made a phyletic evolution study on TAS3 sites, and they found that the duplication events of genes or chromosomes led to the expansion of TAS3 members, though some members were lost in some species. Johnson et al. [20] found some differences between phased small RNA and tasiRNA associated with inflorescence stages in rice. The number of tasiRNA members was very few but they were expressed in nearly every tissue,

while there were hundreds of such small RNA loci distributed in several small RNA clusters that were only expressed at the developmental stage of inflorescence. Moreover, such types of small RNA only existed in a 22-nt phased array with the 5' cleavage site either with the length of 21-nt or 24-nt, which was similar to the specific TAS1/TAS2 genes in $Arabidopsis^{[13]}$ and different from the single-hit model of TAS3. The latest findings showed that these miRNA-mediated phasiRNAs may have originated from a common progenitor sequence and would be conserved during the evolution process^[28].

2 Research about eTMs of miRNA

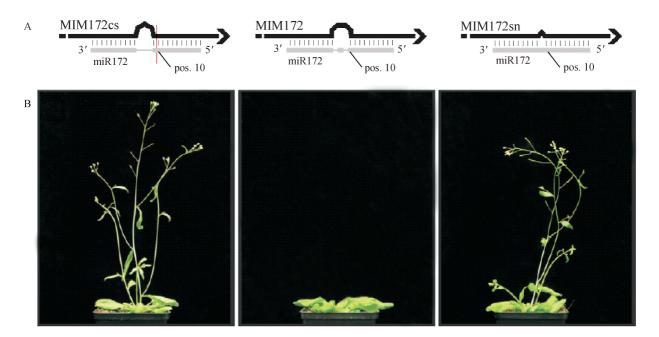
2.1 Definition of eTMs

miRNAs play an important role in the growth and development of animals and plants. In order to realize their regulation functions, miRNAs first bind to the AGO protein to form a complex and specific then target the mRNA complementary base pairing in order to inhibit the expression of the target mRNA[31-32]. Franco-Zorrilla et al. [33] discovered the non-protein coding gene IPS1 (induced by phosphate starvation1) from A. thaliana. The IPS1 gene contained a motif with a sequence complementary to miR399 due to the mismatched loop at the expected miRNA cleavage site, which led to the sequestration of miRNA instead of cleavage. The authentic gene of miR399 is PHO2, which encodes the ubiquitinconjugating enzyme to maintain homeostasis and to perform normal physiological function in the cells[34]. The presence of IPS1 blocks miR399 from targeting PHO2; this type of long non-coding RNA was defined as eTMs.

2.2 Making use of artificial target mimics to analyze the function of miRNA

In plants, many miRNAs have been identified but the majority of their functions were unknown. Discovering the functions of miRNA is increasingly important. In previous research, eTMs showed a negative regulation correlation with its specific miRNA and sequestered the mRNA instead of cleaving it. According to this mechanism, Todesco *et al.* [35] developed some artificial sequences to target mimics in *Arabidopsis* and to forecast the miRNA function. The results showed that 14 functions of all 71 miRNA families were inhibited, resulting in abnormal morphology in *A. thaliana*. Additionally, they found

that the highly conserved miRNA tended to have a great impact on the growth of plants, and the impact of novel miRNAs on plant growth was not obvious. Furthermore, the 3-nt bulge arose between the 10th and 11th bases in the miRNA 5' cleavage site, otherwise the function of target mimics would lose (Fig. 2).



A: Three artificial target mimics. A target mimic with an unmodified central sequence (MIM172cs), which retained complementary to the central portion of miR172 across the cleavage site (red line) opposite position 10 to 11 of the miRNA, did not change flowering time. Modification of the central sequence (TCTA to GAGT; MIM172) restored a three nucleotide bulge found in *IPS1* and generated a functional target mimic, which caused a delay in flowering. However, a single nucleotide mismatch introduced into the center of an authentic miR172 target site (MIM172sn), but without a bulge, was not sufficient to reduce miR172 activity.

Fig. 2 Required bulge at the cleavage site for target mimicry

2.3 Identifying eTMs on a large-scale with bioinformatics methods

The use of artificial target mimics can identify the functions of specific miRNA; however, the identification of eTMs in plants is still at the early stages. The research on the binding rules between miRNA and eTMs has laid the foundation for the identification of eTMs on a large scale. Meng et al. [36] first identified eTMs in the transcriptomes of Arabidopsis TAIR 10 (The Arabidopsis Information Resource, release 10) and rice TIGR 6.1 (The Institute for Genome Research, release 6.1). The results found 300 and 260 potential eTMs in Arabidopsis and rice, respectively. The

following methods were used: First, Ssearch from the FASTA3 program package [downloaded from the FTP site of EBI (European Bioinformatics Institute), ftp://ftp. ebi. ac. uk/pub/software/unix/fasta/][37-38] was used to find the sites in the cDNA sequences that were reverse complementary to the miRNAs. Second, Perl scripts were written to screen the results from Ssearch, obeying the following rules: 1) The 3- to 5-nt bulges must exist within the complementary sites of the cDNAs, and the bulges should be located in the middle of the corresponding miRNAs (definition of the middle positions: 9th to 11th-nt of the 19-nt-long miRNAs, 10th to 11th-nt of the 20-nt ones,

10th to 12th-nt of the 21-nt ones, 11th to 12th-nt of the 22-nt ones, 11th to 13th-nt of the 23-nt ones, and 12th to 13th-nt of the 24-nt ones); 2) for the number of mismatches within the non-middle region of each miRNA, no more than four were allowed, and the consecutive mismatches could not exceed 2 nt; and 3) no bulge was permitted within the non-middle regions of the miRNAs. In addition, they also studied the distribution of the eTMs sites in Arabidopsis and rice and found that most of these sites were in the UTR region while only a small portion was located in the CDS region. GO was applied to do the enrichment analysis for eTMs, and degradome was sequenced to recognize the authentic targets of miRNA, thus the regulatory networks including eTMs, miRNAs, and target genes were successfully constructed.

Banks *et al.* [39] applied computational analysis to identify a numerous endogenous sequences with potential miRNA decoy activity for conserved miRNAs in several plant species. Their data suggested that eTMs can be widespread and may be a component of the global gene expression regulatory network in plants.

Several previous studies identified eTMs through the annotated genome sequences, and Wu et al. [40] developed a new computational method to predict the distribution of eTMs in the intergenic regions and non-coding genes. According to the RNA-Seq data and RT-PCR experiments, they found that more than 40% of the detected predicted eTMs were expressed. They also proved that the eTMs of several miRNA could effectively inhibit the functions of their corresponding miRNAs. Non-coding RNAs, short ORF (open reading frame)-encoding genes (with coding potential of less than 100 amino acids), and intergenic sequences in the Arabidopsis Information Resource (TAIR) 10 for Arabidopsis and the Michigan State University Rice Genome Annotation Project (release 7) annotations were collected and used as the eTMs prediction libraries. eTMs for the 20 selected miRNAs were predicted using local scripts with the following rules: 1) Bulges were only permitted at the 59 end from the 9th to 12th positions of the miRNA sequence; 2) the bulge in eTMs should be composed of only three nucleotides; 3) a perfect nucleotide pairing was required at the 59 end from the 2nd to 8th positions of the miRNA sequence; and 4) except for the central bulge, the total mismatches and G/U pairs within the eTMs and miRNA pairing regions should be no more than three. The distance between qualified intergenic eTMs and their upstream/downstream genes should be longer than 200 nucleotides. As a result, 25 of 36 detected eTMs and 94 of 186 detected eTMs were indeed transcribed in Arabidopsis and rice, respectively. The eTMs for an miRNA can be more than one, with the sequences conservation existing in the target sites.

3 Conclusions

This review describes the recent progress in phasiRNA and eTMs, which are closely related to miRNA. The existence of phasiRNA and eTMs is important in maintaining the homeostasis of plant cells; currently, only a small portion of these was proved to be functional, and more resources should be used to clarify the many mysteries of phasiRNA and eTMs. Fortunately, the bioinformatics was developed quickly and was crucial in identification of phasiRNA and eTMs. development of this sequencing technique and its continuous production of high-throughput data have laid a good foundation for future analysis and research. In addition, the research about small RNA in plants is developing rapidly and occupies a very important area in the plant kingdom. However, the mechanisms of its formation and evolution still remain to be solved.

Acknowledgements We thank Monica (Qiu Mingna) for her careful English editing.

References:

[1] Howell M D, Fahlgren N, Chapman E J, et al. Genome-wide analysis of the RNA-dependent RNA polymerase6/dicer-like4 pathway in *Arabidopsis* reveals dependency on miRNA-and

- tasiRNA-directed targeting. *The Plant Cell*, 2007, 19(3): 926-942.
- [2] Chen H M, Li Y H, Wu S H. Bioinformatic prediction and experimental validation of a microRNA-directed tandem transacting siRNA cascade in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America, 2007, 104(9):3318-3323.
- [3] Zhang C Q, Wang J, Hua X, et al. A mutation degree model for the identification of transcriptional regulatory elements.

 BMC Bioin formatics, 2011, 12:262.
- [4] Rajagopalan R, Vaucheret H, Trejo J, et al. A diverse and evolutionarily fluid set of microRNAs in Arabidopsis thaliana. Genes & Development, 2006, 20(24):3407-3425.
- [5] Liu Y, Wang Y, Zhu Q H, et al. Identification of phasiRNAs in wild rice (Oryza rufipogon). Plant Signaling & Behavior, 2013,8(8):e25079.
- [6] Zhai J, Jeong D H, de Paoli E, et al. MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. Genes & Development, 2011, 25(23):2540-2553.
- [7] Vazquez F, Vaucheret H, Rajagopalan R, et al. Endogenous trans-acting siRNAs regulate the accumulation of Arabidopsis mRNAs. Molecular Cell, 2004, 16(1):69-79.
- [8] Peragine A, Yoshikawa M, Wu G, et al. SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in Arabidopsis. Genes & Development, 2004, 18(19):2368-2379.
- [9] Yoshikawa M, Peragine A, Park M Y, et al. A pathway for the biogenesis of trans-acting siRNAs in Arabidopsis. Genes & Development, 2005, 19(18):2164-2175.
- [10] Axtell M J, Jan C, Rajagopalan R, et al. A two-hit trigger for siRNA biogenesis in plants. Cell, 2006, 127 (3): 565-577.
- [11] Fahlgren N, Montgomery T A, Howell M D, et al.
 Regulation of AUXIN RESPONSE FACTOR3 by TAS3 tasiRNA affects developmental timing and patterning in
 Arabidopsis. Current Biology, 2006,16(9):939-944.
- [12] Williams L, CarlesC C, Osmont K S, et al. A database analysis method identifies an endogenous trans-acting short-interfering RNA that targets the Arabidopsis ARF2, ARF3, and ARF4 genes. Proceedings of the National Academy of Sciences of the United States of America, 2005, 102 (27): 9703-9708
- [13] Allen E, Xie Z, Gustafson A M, et al. microRNA-directed phasing during trans-acting siRNA biogenesis in plants. Cell, 2005, 121(2):207-221.
- [14] Axtell M J, Snyder J A, Bartel D P. Common functions for diverse small RNAs of land plants. The Plant Cell, 2007, 19 (6):1750-1769.
- [15] Zhao T, Li G L, Mi S J, et al. A complex system of small RNAs in the unicellular green alga Chlamydomonas reinhardtii. Genes & Development, 2007, 21(10):1190-1203.

- [16] Heisel S E, Zhang Y, Allen E, et al. Characterization of unique small RNA populations from rice grain. PLoS ONE, 2008, 3(8):e2871.
- [17] Zhu Q H, Spriggs A, Matthew L, et al. A diverse set of microRNAs and microRNA-like small RNAs in developing rice grains. Genome Research, 2008, 18(9):1456-1465.
- [18] Allen E, Xie Z, Gustafson A M, et al. Evolution of microRNA genes by inverted duplication of target gene sequences in Arabidopsis thaliana. Nature Genetics, 2004, 36(12):1282-1290.
- [19] Arif M A, FatTASh I, Ma Z R, et al. Dicer-like3 activity in physcomitrella patens Dicer-like4 mutants causes severe developmental dysfunction and sterility. Molecular Plant, 2012,5(6):1281-1294.
- [20] Johnson C, Kasprzewska A, Tennessen K, et al. Clusters and superclusters of phased small RNAs in the developing inflorescence of rice. Genome Research, 2009, 19 (8): 1429-1440.
- [21] Shivaprasad P V, Chen H M, Patel K, et al. A microRNA superfamily regulates nucleotide binding site: leucine-rich repeats and other mRNAs. The Plant Cell, 2012, 24(3): 859-874
- [22] Arenas-Huertero C, Pérez B, Rabanal F, et al. Conserved and novel miRNAs in the legume Phaseolus vulgaris in response to stress. Plant Molecular Biology, 2009, 70(4): 385-401.
- [23] Xia R, Zhu H, An Y Q, et al. Apple miRNAs and tasiRNAs with novel regulatory networks. Genome Biology, 2012, 13 (6):R47.
- [24] Källman T, Chen J, Gyllenstrand N, et al. A significant fraction of 21-nucleotide small RNA originates from phased degradation of resistance genes in several perennial species.

 Plant Physiology, 2013,162(2):741-754.
- [25] Windels D, Vazquez F. miR393: integrator of environmental cues in auxin signaling? *Plant signaling & Behavior*, 2011, 6 (11):1672-1675.
- [26] Wang Y, Shen D, Bo S P, et al. Sequence variation and selection of small RNAs in domesticated rice. BMC Evolutionary Biology, 2010, 10(1):119.
- [27] Li F, Pignatta D, Bendix C, et al. MicroRNA regulation of plant innate immune receptors. Proceedings of the National Academy of Sciences of the United States of America, 2012, 109(5):1790-1795.
- [28] Xia R, Meyers B C, Liu Z, et al. MicroRNA superfamilies descended from miR390 and their roles in secondary small interfering RNA biogenesis in eudicots. The Plant Cell, 2013, 25(5):1555-1572.
- [29] Song X W, Li P C, Zhai J X, et al. Roles of DCL4 and DCL3b in rice phased small RNA biogenesis. The Plant Journal, 2012,69(3):462-474.
- [30] Shen D, Wang S, Chen H, et al. Molecular phylogeny of miR390-guided trans-acting siRNA genes (TAS3) in the

- grass family. Plant Systematics and Evolution, 2009, 283(1/2):125-132.
- [31] Garcia D. A miracle in plant development: role of microRNAs in cell differentiation and patterning. Seminars in Cell & Developmental Biology, 2008, 19(6):586-595.
- [32] Shukla L I, Chinnusamy V, Sunkar R. The role of microRNAs and other endogenous small RNAs in plant stress responses. Biochimica et Biophysica Acta: Gene Regulatory Mechanisms, 2008, 1779(11):743-748.
- [33] Franco-Zorrilla J M, Valli A, Todesco M, et al. Target mimicry provides a new mechanism for regulation of microRNA activity. Nature Genetics, 2007, 39 (8): 1033-1037.
- [34] Chitwood D H, Timmermans M C P. Target mimics modulate miRNAs. *Nature Genetics*, 2007, 39(8):935-936.
- [35] Todesco M, Rubio-Somoza I, Paz-Ares J, et al. A collection of target mimics for comprehensive analysis of microRNA

- function in Arabidopsis thaliana. PLoS Genetics, 2010, 6 (7):e1001031.
- [36] Meng Y J, Shao C G, Wang H Z, et al. Target mimics: an embedded layer of microRNA-involved gene regulatory networks in plants. BMC Genomics, 2012, 13(1):197.
- [37] Pearson W R. Searching protein sequence libraries: comparison of the sensitivity and selectivity of the Smith-Waterman and FASTA algorithms. *Genomics*, 1991,11(3):635-650.
- [38] Pearson W R. Flexible sequence similarity searching with the FASTA3 program package//Bioinformatics Methods and Protocols. New York, USA: Humana Press, 1999:185-219.
- [39] Banks I R, Zhang Y, Wiggins B E, et al. RNA decoys: an emerging component of plant regulatory networks? Plant Signaling & Behavior, 2012,7(9):1188-1193.
- [40] Wu H J, Wang Z M, Wang M, et al. Widespread long noncoding RNAs as endogenous target mimics for microRNAs in plants. Plant Physiology, 2013, 161(4):1875-1884.

欢迎订阅《浙江大学学报(农业与生命科学版)》

《浙江大学学报(农业与生命科学版)》是浙江大学主办的全国性、综合性、学术性期刊.主要刊登农业 基础科学和应用科学研究论文.内容涵盖农业和生命科学各个领域,包括作物遗传育种;种质资源;植物 保护;生理生态;作物栽培;土壤肥料;园艺;农产品的储藏.保鲜.加工;畜牧.兽医;以及生命科学方面的 专论,研究论文,快讯等.读者对象主要是国内外农业科学研究人员,农业院校的教师和研究生,以及综合 性大学等有关农业科学的研究与管理人员.本刊近年的影响因子、总被引频次两项重要指标均名列全国 农业高校学术期刊的前列;2001年入选中国期刊方阵;2004年教育部全国高校优秀科技期刊评比中荣获 壹等奖,全国优秀农业期刊评比壹等奖;2005年获第三届国家期刊奖百种重点期刊奖;2006年再次荣获 全国农业期刊评比壹等奖,首届中国高校优秀科技期刊奖;荣获中信所2008年度"中国百种杰出学术期 刊"称号;2009年荣获"高校科技期刊优秀编辑质量奖";2010年获"第三届中国高校精品科技期刊奖"; 2011 年获中信所"第二届中国精品科技期刊"和第二届"RCCSE 中国权威学术期刊"称号;2012 年获"第 四届中国高校精品科技期刊奖";现为北京大学图书馆全国综合性农业科学类核心期刊.目前被美国《化 学文摘》(CA),荷兰《文摘与引文数据库》(Scopus),英国《国际农业与生物科学研究中心文摘》(CABI),俄 罗斯《文摘杂志》(AJ),波兰《哥白尼索引》(IC),英国《动物学记录》(ZR),英国《食品科技文摘》(FSTA), 联合国粮农组织农业索引(AGRIS),美国《乌利希期刊指南》(Ulrichsweb)及中国科学引文数据库,《中国 学术期刊文摘》,中国期刊网《中国学术期刊(光盘版)》数据库,万方数据资源系统数字化期刊群数据库等 国际国内重要检索系统收录.

《浙江大学学报(农业与生命科学版)》为双月刊, A4 开本, 120 页, 国内外公开发行. 国内统一刊号: CN 33-1247/S, 国际标准刊号: ISSN 1008-9209, 邮发代号: 32-48, 国外代号: BM4108. 定价每期 25.00 元/人民币, 全年定价 150.00 元/人民币.

《浙江大学学报(农业与生命科学版)》稿约

1 《浙江大学学报(农业与生命科学版)》是浙江大学主办的有关农业和生命科学类学术期刊(双月刊,国内外公开发行),面向国内外征稿.主要刊登农业科学、生物科学与技术、环境科学、食品科学与技术、核农学与技术、遥测遥感技术及其基础学科和相关学科的学术论文与综述.读者对象为国内外相关的科技工作者、高等院校的教师和研究生等.

2 来稿要求及注意事项

- 2.1 来稿务求论点明确,材料翔实,数据准确可靠,论证合理,结论新颖,文字精炼,语句通顺,可用英文或中文写稿.其中,英文文稿需提供中文题名和摘要等;中文文稿必须包括(按顺序):题名(不超过25字),作者姓名及工作单位(单位须至系、所一级,并包括邮编),中文摘要(内容应包括研究的目的、方法、结果和讨论)及关键词(3~8个,尽量使用名词术语的全称),中图分类号,英文题名,作者姓名的汉语拼音(姓大写,名首字母大写),单位译名,邮编,国别,英文摘要(应写出论文主要试验数据和信息,内容需详细,约500个单词;用第三人称撰写,应符合英语语法)和英文关键词,正文,致谢,参考文献(参考文献务须按本刊格式译成英文和中国学术期刊光盘版的要求著录,外文专著尤其要注意注明出版地和出版者).此外,在文章首页页脚注明产生论文的基金项目名称和编号,通信作者和第一作者联系电话及 E-mail 地址.
- 2.2 来稿请登录本刊网站 http://www.journals.zju.edu.cn/agr,点击左侧的"作者投稿",按提示步骤完成投稿流程(第 1 次向本刊投稿的作者需先注册,然后投稿).正文部分用通栏排版,小四号字,务须加宽行距(可采用 1.5 倍行距),并确保校对无误;采用规范化名词术语,量名称及符号(除雷诺数 Re 等 25 个特征数和 pH 值外其他一律用单个字母表示)必须符合国家标准;采用我国法定计量单位(用国际通用符号).尤其要注意必须将亩换算成公顷(hm^2), ppm 换算为 mg/L 或 mg/kg 等.外文字母必须分清大小写和正斜体(注意生物拉丁学名、基因名称和变量用斜体; SI 词头、量单位和函数用正体;矩阵、矢量用黑斜体).
- 2.3 表格编排采用三线表,不带竖线.图、表要少而精,应有足够的信息,具有自明性,即应使读者不查阅正文即可理解图、表的内容.图、表应排在正文首次提及之后.图题、表题、表中横竖项目栏、注释和图中的内容须中、英文对照.数量关系尽量用图表示,文字部分避免罗列与图、表重复的数字.线条图中点线要准确、清晰、美观,不宜过粗或过细.计算机制图请提供制图参数.显微照片和地图,在图中标出比例尺和指北方向,照片要清晰,反差好.
- 2.4 参考文献一律采用英汉对照,引文出处按照在文章中出现的先后顺序编号、排列,即按"顺序编码制"编写.在文中引用处的右上角加方括号标明文献序号.文献作者 3 人以内须全列出,超过 3 人时列前 3 人,后加 et al 或等.英美等国作者均按姓前名后顺序书写,名缩写,不加缩写点.作者应对所引文献的准确性负责,修改时请务必逐篇核查文献的正确性.在参考文献中引用非英文文献时,无论其原文是否用英文形式刊出,均用英文形式列出,并在页码后注明语种.中文期刊文献,根据英文摘要的有无在页码后标注(in Chinese with English abstract)或(in Chinese)等.刊名、书名用斜体.几种主要参考文献的著录格式如下:

【期 刊】 [序号] 作者.题名.刊名,出版年份,卷(期):起止页码.

【专 著】[序号] 作者.书名.版次(初版不写).译者(原著不写).出版地:出版者,出版年:引文页码.

【论文集】[序号] 析出文献作者.题名//主编者.论文集名.出版地:出版者,出版年:引文页码.

【学位论文】 [序号] 作者.题名.保存地点:保存单位,年份.

【专利文献】 [序号] 专利所有者.题名:专利国别,专利号.公告日期.获取和访问路径.

【电子文献】 [序号] 主要责任者. 题名:其他题名信息[文献类型标志/文献载体标志]. 出版地:出版者,出版年. 获取和访问路径.

- *期刊:[1] 项延军,李新芝,王小德.5 种藤本植物的抗寒性研究初探.浙江大学学报:农业与生命科学版,2011,37(4):421-424.

 Xiang Y J, Li X Z, Wang X D. Preliminary study on cold resistance of five different lianas. *Journal of Zhejiang University: Agriculture and Life Sciences*, 2011,37(4):421-424. (in Chinese with English abstract)
- 2.5 来稿如不符上述要求,编辑部均退请作者修改补充后再考虑是否受理.
- 2.6 来稿经修订,拟于刊用,作者须按标准缴纳版面费.为倡导绿色办刊理念,论文正式刊出后不赠送当期样刊,本刊提供含当期封面、目录的 PDF 单篇文档,若需纸质样刊需提前订购.为使您的研究成果得到更广泛的传播与交流,您可以给编辑部提供一份名单,我们会把您的全文 PDF 文件发送给这些专家.
- 2.7 来稿文责自负.本刊已加入中国科学引文数据库、中国科技论文统计分析数据库、中国学术期刊综合评价数据库、万方数据资源系统数字化期刊群数据库等.本刊发表的所有文章,均同时在上述载体发表.作者需交寄"版权转让协议".作者著作权使用费与本刊稿酬一次性付给.

《浙江大学学报(农业与生命科学版)》编辑室

封面设计: 万昌平

Zhejiang Daxue Xuebao

(Nongye yu Shengming Kexue Ban)

本刊被下列重要检索系统收录:

美国《化学文摘》(CA)

荷兰《文摘与引文数据库》(Scopus)

英国《国际农业与生物科学研究中心文摘》(CABI)

俄罗斯《文摘杂志》(AJ)

波兰《哥白尼索引》(IC)

英国《动物学记录》(ZR)

英国《食品科技文摘》(FSTA)

联合国粮农组织农业索引(AGRIS)

美国《乌利希期刊指南》(Ulrichsweb)

北京大学图书馆中文核心期刊

中国科技核心期刊

RCCSE中国核心学术期刊

浙江大学学报

(农业与生命科学版) (双月刊 1956年创刊)

第40卷 第4期(总第185期) 2014年7月20日

Journal of Zhejiang University (Agriculture & Life Sciences)

(Bimonthly, Started in 1956)

Vol.40 No.4 (General Serial No. 185) July 20, 2014

主管部门:中华人民共和国教育部 Responsible Unit: Ministry of Education of the People's Republic of China

主办单位 浙江大学 Sponsor Unit: Zhejiang University

主 编:朱军 Editor in Chief: Zhu Jun

编辑单位 《浙江大学学报(农业与生命科学版)》编辑部 Editing Unit: Editorial Board of Journal of Zhejiang University (Agric. & Life Sci.)

通信地址 杭州市天目山路148号 邮编310028 Commun. Add: 148 Tianmushan Road, Hangzhou 310028, China

网络地址Http://www.journals.zju.edu.cn/agr Website: Http://www.journals.zju.edu.cn/agr

出版单位 浙江大学出版社有限责任公司Publishing Unit:Zhejiang University Press Co., Ltd.印刷单位 杭州丰源印刷有限公司Printing Unit:Hangzhou Fengyuan Printing Co., Ltd.

国内发行:浙江省报刊发行局Domestic Distribution:Zhejiang Post Office订购处:全国各地邮局Domestic Subscription:Local Post Office

国外发行: 中国国际图书贸易集团有限公司 Overseas Distribution: China International Book Trading Corporation

(北京市海淀区车公庄西路35号 邮编:100048) (35 Chegongzhuang Xilu, Haidian District, Beijing 100048, China)