



生物信息学方法挖掘小麦中的 microRNAs 及其靶基因

韩友生, 栾福磊, 朱洪亮, 邵毅, 陈安均, 卢承文, 罗云波, 朱本忠*

中国农业大学食品科学与营养科技学院, 北京 100083

* 联系人, E-mail: cauzbz@yahoo.com.cn

收稿日期: 2009-04-09; 接受日期: 2009-04-23

国家自然科学基金(批准号: 30600421 和 30871741)资助项目

摘要 microRNAs(miRNAs)是最近发现的一种内源、非编码、长度约为 21 nt 的小分子 RNA, 在转录后水平起调控基因表达的作用. 先前的研究发现, 植物中的 miRNAs 具有高度的保守性, 这为在其他植物物种中通过同源比对发现保守的 miRNAs 提供了思路. 本文利用 EST 和 GSS 分析法在小麦中预测 miRNA 及其靶基因. 简单过程是, 将其他植物物种中已经发现的 miRNA 与小麦的 EST 和 GSS 数据库比对, 经过一系列的标准筛选、预测小麦 miRNA. 通过这一策略, 本文共得到属于 18 个家族的 37 条小麦 miRNA, 其中有 10 条保守的 miRNA 是第一次被发现. 发现 miR395 是一个非常特殊的家族, 因为它的 3 个成员成簇的存在形式与动物 miRNAs 非常相似. 本文还利用新发现的小麦 miRNAs 通过在线软件 miRU 与编码蛋白的数据库比对, 总共预测到 361 个靶基因, 其中一些靶基因参与到了小麦的生长发育、新陈代谢及抗逆过程.

关键词

表达序列标签
(express sequence tag, EST)
基因组概览序列
(genome survey sequence, GSS)
microRNAs
靶基因
小麦

microRNAs(miRNAs)是一类新发现的、内源、非编码小分子 RNA, 在转录后水平通过抑制翻译或者引起靶 mRNA 降解的方式调控基因表达^[1]. miRNAs 在植物生长发育和其他生理过程中通过调控功能基因的表达而起重要作用^[2]. 植物 miRNAs 的成熟过程涉及到一些关键的酶, 诸如 DCL1^[1], HASTY^[3,4]. 成熟 miRNAs 被整合到一种基因沉默复合物(RISC)中, 该复合物靶定靶 mRNA 而引起靶基因的降解^[1].

随着越来越多的 miRNAs 被发现, 一些物种特异的 miRNAs 的数量也在以惊人的速度被发现^[5-7], 并且在植物领域中很多都是高度保守的^[8]. 目前, 已经有 4 种方法被应用在多种物种中发现新的 miRNAs:

(1) 遗传筛选; (2) 构建小分子 RNA 基因文库和深度测序; (3) 传统的生物信息学; (4) 表达序列标签 (express sequence tag, EST)和基因组概览序列 (genome survey sequence, GSS)分析法. 遗传筛选是最早被用于发现和分析 miRNA 的方法, miRNAs 家族最早的两个成员, *lin-4* 和 *let-7*, 就是通过遗传筛选的方法在线虫(*Caenorhabditis elegans*)中被发现的^[9,10]. 然而, 由于该方法耗时、价格昂贵而只被应用在 miRNAs 研究的初期阶段. 另外一种发现新的小分子 RNA(包括 miRNA 和 siRNA)的方法是构建小分子 RNA 基因文库, 这也是发现 miRNAs 的主要方式^[6,7,11]. 通过这种方法, 在一些模式植物中发现了大

量的 miRNAs, 包括拟南芥(*Arabidopsis thaliana*)^[6]和水稻(*Oryza sativa*)^[7]. 该方法需要在总 RNA 中分离小分子 RNA, 并且需要两个特殊接头, 因此费时费力. 另外, 该方法最大的缺陷是很难发现暂时表达或组织特异性的 miRNAs. 最近发明的深度测序技术是发现 miRNAs 的强有力的工具, 其同时还能对 miRNAs 进行定量分析. 该技术为在一些基因组序列未知的物种中大量发现 miRNAs 提供了可能, 如番茄(*Lycopersicon esculentum*)^[12]. 目前已经开发了 3 种深度测序技术, 454, Solexa 和 Illumina. 利用这些技术已经在一些基因组序列未知的植物物种中发现了大量的新 miRNAs, 如红豆杉(*Taxus chinensis*)^[13]和番茄^[12]. 但该技术比较耗时, 而且需要大量数据分析, 使工作非常繁琐. 与前两种方法相比, 第 3 种方法是一种高通量的、简便的方法, 即依据全基因组序列的计算机搜索. 该方法主要利用一些 miRNAs 及其前体的特性对 miRNAs 进行定性^[10,14]. 尽管有些科研小组已经开发了一些计算机软件用于在拟南芥和水稻等模式植物中发现 miRNAs^[14-16], 但计算机搜索的一个明显缺点是它需要全基因序列, 这大大限制了该方法的应用, 因为只有少数的模式植物的全基因序列已知. 为了验证这些预测的 miRNAs 的表达, 通常还需要 Northern 印记杂交. 第 4 种发现保守 miRNAs 的方法是表达序列标签(EST)分析法, 由 Zhang 等人^[17]根据植物 miRNAs 的高度保守性发明. 简单地说, 利用模式植物, 如拟南芥和水稻中已经发现的 miRNAs 与其他物种的 EST 或/和 GSS 序列进行比对, 通过一系列的筛选标准(如错配的数量、发夹结构等)得到. 利用这种方法, Zhang 等人^[17]在 18694 条 EST 序列中发现了 60 个植物物种的 338 条 miRNAs, 其中有些涉及到植物的生物和非生物胁迫. 最近, 随着越来越多的 miRNAs 被发现, 在更多物种中发现了 miRNAs, 如棉花(*Gossypium hirsutum*)、玉米(*Zea mays*)^[18,19]. 这些结果表明, EST 分析法是在基因组序列未知的物种中发现保守 miRNAs 的强有力的方法.

小麦(*Triticum aestivum* L.)是全球种植量最大的农作物之一, 是亚洲、欧洲、美洲等地人们的主要食物来源, 也是生产淀粉、蛋白质和生物能源的重要资源. 目前, 有 4 篇文章分别通过 EST 分析、构建小分

子 RNA 文库直接克隆和数据挖掘的方法在小麦中报道了 miRNA^[17,20-22]. 由于在 2005 年发现的植物 miRNA 很少, 并且主要是拟南芥和水稻的 miRNA, Zhang 等人^[17]利用 EST 分析的方法只发现了属于 9 个家族的 19 条保守的 miRNA. 最近, Yao 等人^[21]通过构建小分子 RNA 文库的方法克隆得到了属于 43 个家族的 58 条 miRNA, 其中 23 条为小麦特有. Dryanova 等人^[22]通过数据分析挖掘的方法也得到了一些小麦 miRNAs. 但是, 这些方法仍没有将小麦中的 miRNAs 充分挖掘. 本研究利用所有被 miRNA 数据库 miRBase 收录的植物 miRNA(<http://miRNA.sanger.ac.uk>)(Release 10.0, November 2007)^[23]为搜索序列与小麦的 EST 和基因组概览序列(genome survey sequence, GSS)数据库比对, 总共发现了 37 条小麦 miRNA, 包括先前报道的 27 条, 另外的 10 条为首次发现. 本研究还利用发现的 miRNA 预测到 361 个潜在的靶基因, 一部分为转录因子, 还有一些是在小麦生长发育和抗逆过程中起重要调控作用的蛋白.

1 材料与方法

1.1 miRNA 的搜索序列

在 miRNA 的数据库 miRBase(<http://microrna.sanger.ac.uk/cgi-bin/Ram/mirna/browse.pl>; Release 10.0, November 2007)^[23]中下载所有已知植物的 miRNA 及其前体序列, 包括 184 条拟南芥 miRNA、243 条水稻 miRNA、220 条小立碗藓(*Physcomitrella patens*) miRNA、215 条毛果杨(*Populus trichocarpa*) miRNA 和 95 条玉米 miRNA, 剩余的 miRNA 来自于其他模式植物, 如大豆(*Glycine max*)、苜蓿(*Medicago truncatula*)、甘蔗(*Saccharum officinarum*)、高粱(*Sorghum bicolor*)等. 为避免 miRNA 的重复搜索, 本研究首先删除了属于同一个家族并且具有相同序列的 miRNA, 剩余的 miRNA 序列被当作预测小麦 miRNA 的搜索序列.

1.2 小麦的表达序列标签(ESTs)、基因组概览序列(GSSs)和小麦的 mRNA 序列

小麦的 EST 和 GSS 序列在日本的 DNA 数据库 (DDBJ, version 2.2.15)(<http://www.ddbj.nig.ac.jp>)中下

载得到. 用于预测小麦 miRNA 靶基因的小麦 mRNA 则在 TIGR 基因组数据库(<http://tigr.org>)中下载得到.

1.3 潜在小麦 miRNA 的 EST 和 GSS 分析法预测

图 1 为小麦 miRNA 的搜索过程. 简单地说, 对于在拟南芥、水稻等植物中得到的 miRNA 进行相互比对, 去除掉重复序列. 将剩余的不重复的、序列唯一的 miRNA 与小麦的 EST 和 GSS 数据库进行 BLAST(<http://blast.ddbj.nig.ac.jp/top-e.html>), 与搜索 miRNA 存在不多于 4 个碱基错配的小麦 EST 和 GSS 序列筛选出来, 得到候选的小麦 miRNA 序列. 将候选的 EST 和 GSS 序列进行一次自身比对, 去除掉具有相同序列的 miRNA; 然后, 将符合上述条件的候选 EST 和 GSS 序列与 Rfam 数据库比对, 去除掉编码 rRNA, tRNA 等的序列; 最后, 将候选 EST 和 GSS 序列与 DDBJ 收录的编码蛋白质的 RNA 序列进行一次 BLASTX 比对, 去除掉编码蛋白质的 EST 和 GSS; 剩下的序列用在线软件 MFOLD([\[for.rpi.edu/applications/mfold/old/rna\]\(http://for.rpi.edu/applications/mfold/old/rna\)\)^{\[24\]}进行二级结构折叠分析, 得到具有发夹结构的 miRNA. 用于区分小麦 miRNA 与其他小分子 RNA 的标准如下^{\[17,25\]}: \(1\) 考虑到 miRNA 在植物物种间的进化, 允许小麦 miRNA 与搜索 miRNA 序列之间有不超 4 个的错配碱基; \(2\) miRNA 的前体能折叠成发夹结构的二级结构, 并且成熟的 miRNA 位于发夹结构的一条臂上; \(3\) miRNA 前体的二级结构具有较高的绝对值最小折叠自由能\(negative minimal free energies, MFEs\)和较高的最小折叠自由能系数\(negative minimal free energy index, MFEI\); \(4\) miRNA 前体中 A+U 的含量在 30%~70%之间; \(5\) 成熟的 miRNA 序列与另一条臂上的互补序列之间的错配碱基不超过 6 个; \(6\) 成熟的 miRNA 序列与另一条臂上的互补序列之间不允许存在环或者缺口. 成熟的小麦 miRNA 与其他植物的 miRNA 用 DNAMAN 软件包\(Lynnon corporation, Quebec, Canada J3V, 9M5\)比对, 系统发生树用 DNAMAN 软件包构建. 属于同一家族但来自不同植](http://www.bioin-</p>
</div>
<div data-bbox=)

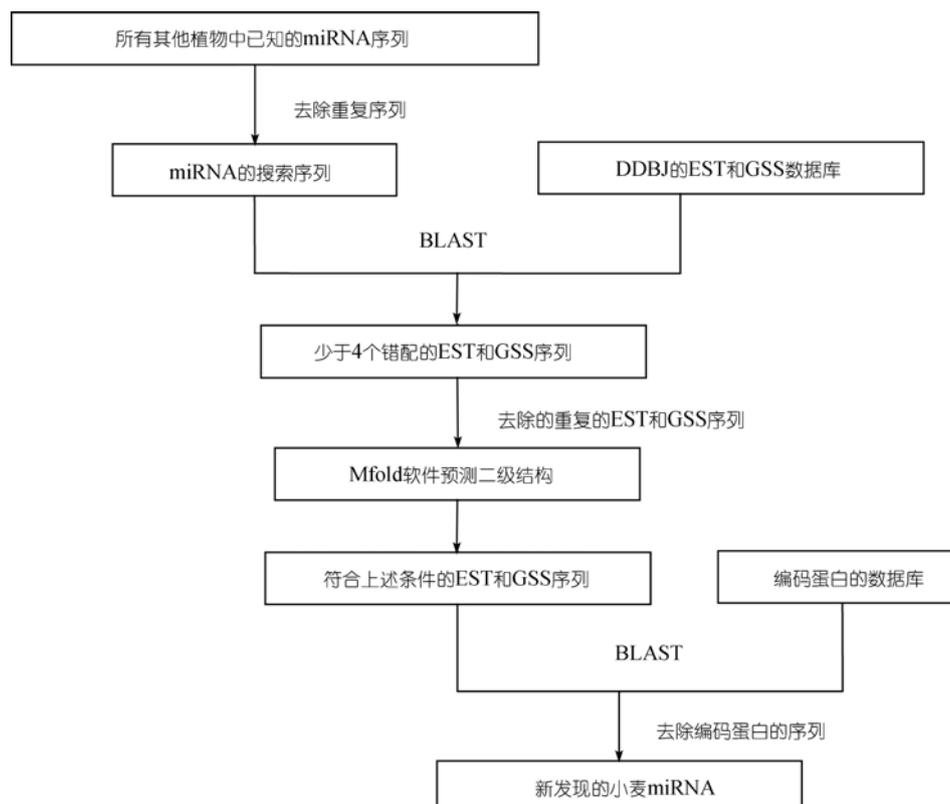


图 1 小麦 miRNA 的预测分析过程

物的 miRNA 的前体用在线软件 MULTALIN (<http://prodes.toulouse.inra.fr/multalin/multalin.html>)^[26]进行多序列比对。

1.4 小麦 miRNA 靶基因的预测

动物中, miRNA 通过与靶 mRNA 部分互补而抑制 mRNA 的翻译. 与动物 miRNA 的作用方式不同, 植物 miRNA 主要通过完全互补或者近乎完全互补的方式识别靶 mRNA, 并引起靶 mRNA 的降解^[1]. 利用这一策略就很容易通过同源性搜索在植物中预测 miRNA 的靶基因. 目前, 有相当一部分预测的靶基因已经得到实验方法的验证. Zhang 等人^[17]利用在线软件 miRU(<http://bioinfo3.noble.org/miRU.htm>)在各种植物中预测 miRNA 靶基因, 取得了较好的结果. 本研究利用 miRU 在 TIGR 基因组数据库(<http://www.tigr.org/>)预测通过上述方法发现的小麦 miRNA 的靶基因. 成熟的 miRNA 与其靶基因 mRNA 之间只允许不超过 4 个错配碱基, 并且不允许存在缺口.

2 结果与讨论

2.1 先前已经报道的小麦 miRNA

为了发现小麦中的保守的 miRNA, 本研究利用 Zhang 等人^[17]发明的 EST 和 GSS 分析方法, 根据 miRNA 在植物间的保守性, 预测小麦的 miRNA 并采用这种方法研究其在植物物种间的保守性. 在小麦中共发现了属于 18 个家族的 37 条 miRNA, 其中包括已经被发现的 27 条 miRNA^[17,20-22], 和 10 条在本研究中首次报道的小麦 miRNA. 有趣的是, 在所有的 37 条发现的小麦 miRNA 中, 只有 3 条是在 GSS 数据库中得到的, 分别是 miR319a, miR169*和 miR166, 其余的 34 条则全部在 EST 数据库中发现. 通过构建小分子 RNA 文库和筛选, Yao 等人^[21]克隆到了 58 个小麦 miRNA, 其中 35 条属于保守的 20 个 miRNA 家族, 剩余的 miRNA 为小麦所特有. 本研究通过保守性分析, 发现了先前已发现的属于 13 个家族的 27 条小麦 miRNA(表 1), 但每个家族的成员数量有所不同. 因为先前实验已经克隆得到这些 miRNA, 本研究未对这些 miRNA 做特性分析. 虽然 Yao 等人^[21]在小麦的小分子 RNA 文库中克隆得到了 miR165, miR168, miR390, miR393 和 miR479, 但本研究通过 EST 和

GSS 分析没有预测到这些 miRNA, 可能是因为小麦中 EST 和 GSS 的数量有限.

2.2 最新通过 EST 和 GSS 分析发现的小麦 miRNA

本研究利用其他植物, 如拟南芥、水稻、玉米中的 miRNA 与小麦 EST 和 GSS 数据库比对, 除了上述已报道的 27 条 miRNA 外, 还发现了其他属于 4 个家族的 10 条新发现的小麦 miRNA(表 2). 为了将新发现的 miRNA 与其他的小分子 RNA 如 siRNA 区分开来, 需要对这些小分子 RNA 进行特性分析, 如 miRNA 前体(pre-miRNA)中 A+U 含量、pre-miRNA 长度、前体二级结构、最小折叠自由能(MFE)和最小折叠自由能指数(MFEI), 其中 MFEI 是将 miRNA 与其他小分子 RNA 区分开来的重要指标.

经过 MFOLD 在线软件折叠, 发现新发现的小麦 miRNA 的前体都能折叠成典型的发夹结构(图 2 和 3). 与动物 miRNA 的前体长度变化不大(一般为 70~90 nt)的现象相比, 小麦 miRNA 的前体的长度变化很大, 在 49~252 nt 的范围内(图 4), 这与其他植物中的 miRNA 前体长度一致. 先前报道证明了 miRNA 的前体中 A+U 的含量高于 G+C 的含量以使前体更加稳定^[27]. 本研究中心发现的小麦 miRNA 前体中, A+U 的平均含量也高于 G+C 的平均含量, 平均含量为 58.33%. 有趣的是, 成熟的 miRNA 在前体茎环结构的任何一条臂上的几率是相同的. 但在本研究中, 新发现的 10 条小麦 miRNA 中只有 miR395a 和 miR395c 在其前体的 5'端, 剩下的 8 条都定位在其前体的 3'端. 另外, 即使综合考虑发现的全部 37 条小麦 miRNA, 也只有 10 条定位在其前体 5'端. 目前还不清楚这一现象是不是小麦 miRNA 所特有. 另外一个将 miRNA 与其他小分子 RNA 区分开来的重要特性是 MFEI. 先前报道^[25]的 miRNA 前体的 MFEI 的平均值在 1.0 左右. 本研究新发现的小麦 miRNA 的 MFEI 的平均值是 1.11, 与先前报道的结果一致, 比其他小分子 RNA 的 MFEI 值高, 如 rRNAs (0.59), tRNAs(0.64)^[25]. 所有的这些特性表明 10 条新发现的小分子 RNA 极有可能是小麦 miRNA.

新发现的小麦 miRNA 家族成员的分布情况与其他植物物种相似. 最大的家族是 miR818, 有 5 个成员; miR395 有 4 个成员; miR156 和 miR167 各有 3 个成

表 1 与先前的发现相同的小麦 miRNAs

miRNAs	序列(5'—3')	EST 或 GSS	LP/nt	定位	(A+U)%	MFEs	MFEIs
miR156a	UGACAGAAGAGAG <u>A</u> GAGCAC	CD454302	306	3'	46.7	115.6	0.91
miR156b	UGACAGAAGAGAG <u>A</u> GAGCAU	BJ319553	52	5'	51.92	42.30	0.85
miR156c	<u>C</u> GACAGAGGAAAGAGAGCAC	CK205751	42	3'	59.52	31.20	0.87
miR159a	UUUGGAUUUGAAGGGAGCUCU <u>G</u>	CJ898977	173	3'	46.82	80.70	0.99
miR159b	UUUGGAUUUGAAGGGAGCUCU <u>G</u>	CA484819	252	3'	53.57	103.2	0.87
miR160	UGCCUGGCUCCUGUAUGCCA	CJ641547	50	5'	66.00	51.80	1.35
miR164a	UGGAGAAGCAGGGCACGUGCA	CA704421	142	5'	33.09	92.50	0.97
miR164b	UGGAGAAGCAGG <u>U</u> CACGUGCG	CD899685	60	5'	35.89	25.00	0.84
miR166	CCGGACCAGGCUUCAAU <u>U</u> UGC	EI674180	58	3'	44.83	31.30	0.86
miR167a	UGAAGCUGCCAGCAUGAUCUA	CK209908	60	5'	48.33	28.70	0.92
miR167b	UGAAGCUGCCAGCAUGAUCUA	CJ833771	90	5'	46.67	35.90	0.91
miR167c	UGAAGCUG <u>A</u> CAGCAUGAUCUA	CK209889	60	5'	50.00	33.30	1.11
miR169*	<u>U</u> U <u>C</u> A <u>G</u> CAAGUUGACCUUGG <u>C</u> A	EI673529	46	3'	43.47	41.20	1.29
miR171	UGAUUGAGCCG <u>U</u> GCCAAUAUC	CD910903	94	3'	37.24	44.98	0.87
miR172a	AGAAUCCUGAUGAUGCUGCAA	CA694228	46	3'	43.47	25.80	0.96
miR172b	AGAAUCCUGAUGAUGCUG <u>C</u> U	CJ942609	90	3'	38.89	32.20	0.85
miR319a	UUGGACUGAAGGG <u>U</u> GCUCCCG	CV774688	47	3'	39.78	22.80	0.79
miR319b	UUGGACUGAAGGGAGCUC <u>C</u> U	CZ889526	198	3'	48.99	97.40	0.94
miR395	<u>G</u> UGAAGUGUUUGGGGGAACUC	CV763592	66	3'	54.55	34.90	1.16
miR396	UCCACAGGCUUUCUUGAACUG	CJ776495	137	5'	45.26	59.0344	0.78
miR398	UGUGUUCUCAGGUC <u>G</u> CCCCCG	CJ711035	86	3'	44.19	0.30	0.93
miR399a	UGCCAAAGGAGAGUUGCC <u>C</u> UG	CJ667854	79	3'	41.77	44.90	1.13
miR399b	UGCCAAAGGAGAA <u>U</u> UGCC <u>C</u> UG	CJ666653	114	3'	40.35	58.18	0.85
miR408a	<u>C</u> UGCACUGCCUCU <u>U</u> CCCGGC	BE419354	113	3'	47.79	40.23	0.92
miR408b	<u>C</u> UGCACUGCCUCUG <u>C</u> CCCGGC	CD907090	115	3'	44.35	40.56	0.88
miR444a	UUGCUGCCUCAAGCUUGCUGC	CK200584	151	3'	50.33	72.69	0.96
miR444b	UUGCUG <u>U</u> CUCAAGCUUGC <u>A</u>	DR738640	113	3'	54.87	56.10	1.10

LP: 前体长度; MFE: 最小折叠自由能; MFEI: 最小折叠自由能指数

表 2 新发现的小麦 miRNAs 及其特性分析

miRNAs	序列(5'—3')	EST	LP/nt	定位	(A+U)%	MFEs	MFEIs
miR395a	CUGAAGUGUUUGGGGGAACUC	CK194045	162	5'	56.25	57.31	0.81
miR395b	<u>A</u> UGAAGUGUUUGGGGGAACUC	CK194045	252	3'	57.94	96.61	0.91
miR395c	<u>G</u> UGAAGUGU <u>U</u> UGGGGGAACUC	CK194045	123	5'	50.41	49.60	0.83
miR414	UCAUC <u>A</u> UCAUCAUCGUCA	CJ967189	49	3'	61.22	13.60	0.72
miR818a	AAC <u>G</u> UCUUAUAUUAUGGGACGG	EB513609	69	3'	68.12	27.10	1.23
miR818b	ACA <u>U</u> UCUUAUAUUAUGGGACGG	DR739383	72	3'	63.89	28.50	1.10
miR818c	A <u>U</u> GUUCUUAUAUUAUGGGACGG	AL816538	69	3'	63.77	45.80	1.83
miR818d	AC <u>G</u> CUUAUAUUAUGGGACGG	CN010552	69	3'	62.33	41.40	1.59
miR818e	AAC <u>G</u> UCUUAUAUUAUGGGACGG	CK210390	68	3'	64.71	33.70	1.40
miR835	UUCUUG <u>C</u> AUAUGUUCUUA <u>A</u> C	CK171746	124	3'	34.68	58.40	0.72

LP: 前体长度; MFE: 最小折叠自由能; MFEI: 最小折叠自由能指数

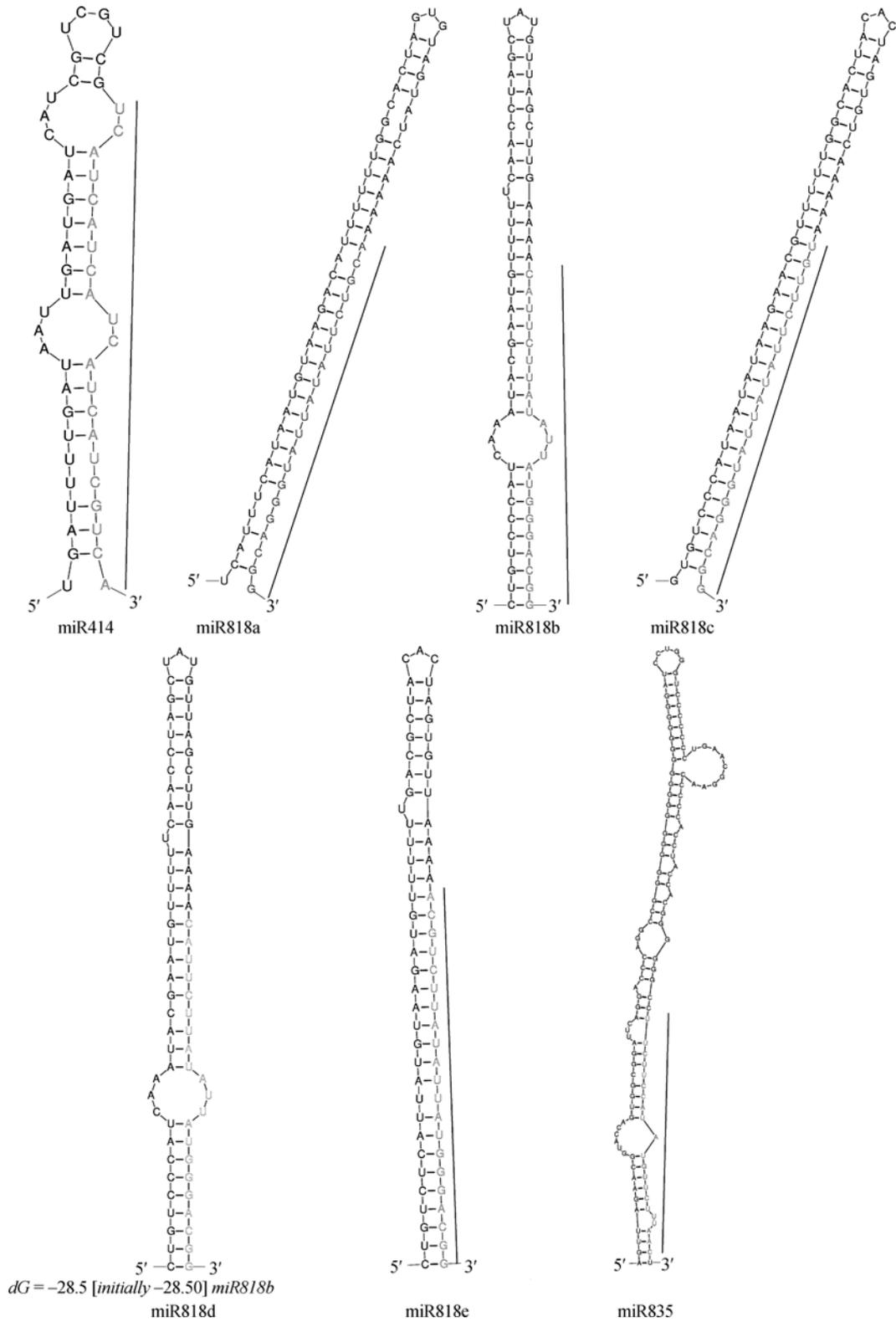


图 2 新发现的小麦 miRNA 二级结构

下画线所示成熟 miRNA 序列

员; miR159, miR164, miR172, miR319, miR399, miR408 和 miR444 各自有两个成员; 其余家族都只有一个成员. 这可能预示着 miR818 和 miR395 在小麦的生长发育和抗逆过程中起着重要作用.

2.3 成簇存在的 3 个 miR395 家族成员

动物 miRNA 的一个普遍特性是属于同一个家族的几条 miRNA 通常成簇的聚集在一起, 然后一起转录成一条多顺反子 mRNA, 一起被剪切成成熟的 miRNA [28]. 但在植物中这种成簇存在的 miRNA 只有少数个例报道 [2,8,29,30]. miR395 是一个在拟南芥和水稻中发现的成簇存在的植物 miRNA [2,11]. 拟南芥中 miR395 的 3 个成员成簇的存在于同一条多顺反子中. 就这一点而言, miR395 与动物 miRNA 很相似. 另外一个成簇存在的植物 miRNA 是在藓类植物中发现的 miR1219(miR1219a 和 miR1219b 由大约 200 bp 的片

段隔开) [30]. 最近, 杨树 miR950 家族的两个成员 (miR950a 和 miR950b) 也是成簇存在的 [29]. 本研究也发现在小麦中 miR395 家族的 3 个成员 (miR395a, miR395b 和 miR395c) 在一条小麦 EST 中成簇存在 (CK194045, 图 3). 有趣的是, miR395a 和 miR395b 拥有相同的前体, 但是方向却相反, miR395c 自己单独拥有一个前体. 水稻 miR395 的 3 个成员却不存在这种情况. 这种 miR395 家族在组织方面的复杂性可能预示着 miRNA 在加工成熟方面的复杂性.

进一步, 本研究分析了小麦的 miR395 和其他植物的 miR395 的保守性. 结果显示, 成熟 miR395 区域的保守性要远远高于其前体的其他部分. 虽然, miR395 成熟序列的个别位点会发生突变, 如 5' 端最容易发生突变 (图 5), 但更多的位点是保守的. 这些高度保守的序列段对于 miRNA 识别其靶位点很重要.

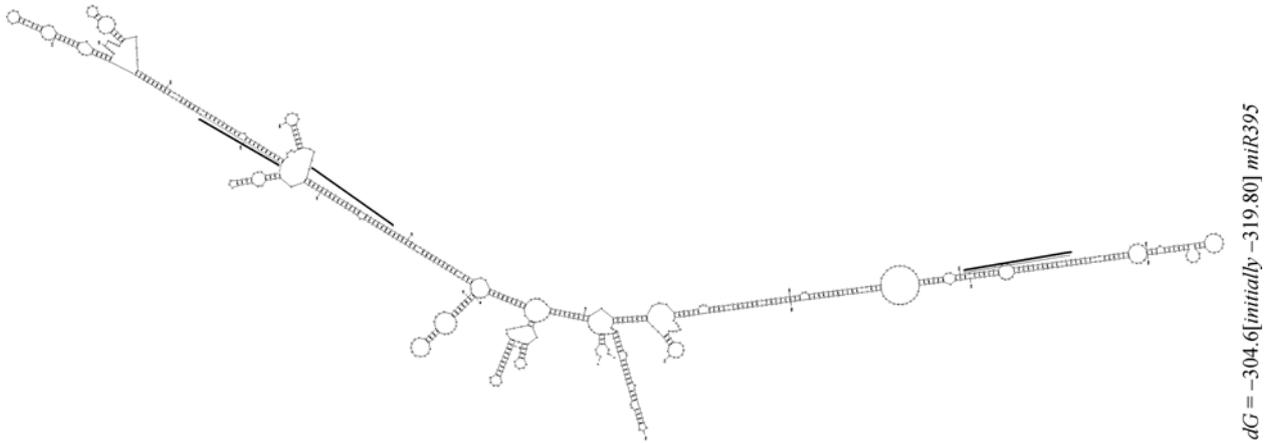


图 3 小麦 miR395 家族的 3 个成员, miR395a, miR395b 和 miR395c, 成簇存在

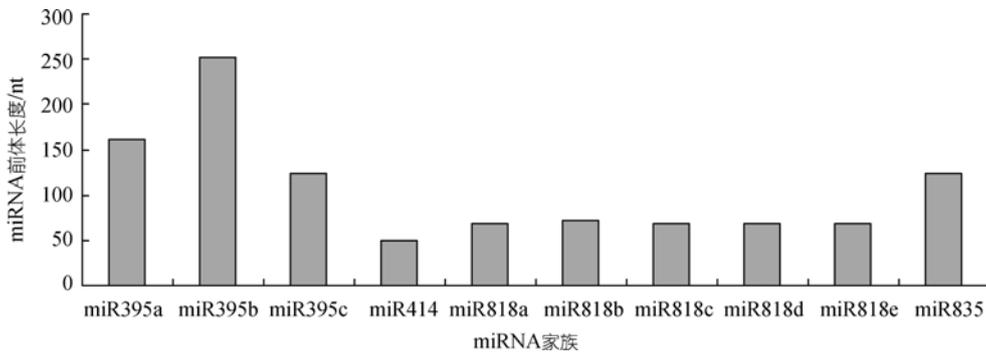


图 4 新发现小麦 miRNA 前体长度变化

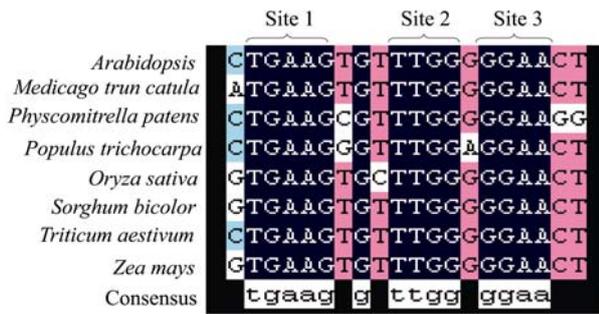


图 5 小麦 miR395a 与其他植物物种中的同源体序列比对

2.4 小麦 miRNA 的潜在靶基因

先前的研究已经成功地预测了一些保守 miRNA 家族的靶基因, 在一些模式植物尤其是拟南芥中通

过 5'RACE 实验的方法验证了靶基因的正确性, 并初步探明了这些 miRNA 在植物生长发育过程中的功能. 随后的实验证明这些 miRNA 的功能涉及到植物组织器官发育, 如花组织极性、叶的多态性和根的生长、生物和非生物胁迫、信号转导等^[16,31-34].

本研究利用上述发现的小麦 miRNA 通过在线软件 miRU 在小麦的 mRNA 数据库中搜索小麦 miRNA 的靶基因, 共找到 37 条小麦 miRNA 的 361 个潜在靶基因. 与以前的研究结果相似, 预测的小麦 miRNA 的 361 个靶基因涉及小麦的生长发育、信号转导和胁迫, 并且 miRNA 倾向于靶定转录因子(表 3). 另外一个大的靶基因家族涉及到小麦的胁迫反应, 这些靶

表 3 小麦 miRNAs 的主要靶基因

miRNAs	靶定蛋白	功能	靶定基因
miR159	MYB33	转录因子	TC194473
	细胞周期相关 Mob1 类蛋白	生长发育	TC224747
	富含亮氨酸的核苷酸结合蛋白 1	新陈代谢	TC229221
	1 类拓扑异构酶	新陈代谢	TC226217
miR160	ARF10	转录因子	TC220268, TC223417
	UDP-葡萄糖基转移酶	新陈代谢	TC215559
miR164	NAC1	转录因子	TC224059, TC224062
	NAC2	转录因子	TC224408
	MAP 激酶 2	新陈代谢	CA681504, TC203027
	同源亮氨酸锌指拉链蛋白	转录因子	TC213254, TC214742
miR167	ARF8	转录因子	TC195542
miR169	CAS	新陈代谢	TC215924
miR172	ORF2	转录因子	CA424651
	APETALA2-类蛋白	转录因子	TC221786, TC209305, TC209306
miR319	PCF8	转录因子	TC199472
	PCF6	转录因子	TC230286
miR395	SAC 结构域蛋白 2	转录因子	TC220251
	Poly(A)聚合酶	新陈代谢	TC189870
	线粒体 ATP 硫化酶	新陈代谢	TC191731
miR396	生长调节因子	转录因子	CK209519
miR414	核小体组装蛋白	新陈代谢	TC191077
	A 大亚体(TFIIA-L1)	转录因子	TC194829
	DNA-结合类蛋白	转录因子	CD891478
	细胞伸长调节蛋白	生长发育	CK158764
	合算聚腺苷化 RNA 结合蛋白 NAB3	生长发育	TC189588
	ATP 酶亚体 6	新陈代谢	TC219270
miR818	抗病相关蛋白 4	胁迫响应	CD936868, TC213259

基因可能会在很大程度上影响小麦的产量和质量。有趣的是, 本研究发现 miR414 在小麦中有多达 120 个靶基因, 尽管其中一部分基因编码的是功能未知的蛋白。

2.5 小麦 miRNA 的靶基因的多样性和主要功能分析

植物 miRNA 靶基因的一个特点是倾向于靶定转录因子^[35]。本研究预测得到的小麦 miRNA 涉及到小麦的生长发育、对环境胁迫的响应。其中一部分靶基因负责编码转录因子, 这些转录因子调节着小麦生长发育和其他的生理代谢过程。例如, 小麦 miR159 (最初叫作 miR-JAW) 预测的靶基因是 MYB3 转录因子, MYB3 是控制叶子正常发育的调控因子。通过在拟南芥中过量表达对 miR159 有抗性的 MYB3, 转基因拟南芥呈现出卷曲的叶片^[34], 表明 miR159 对 MYB3 的负调控对植物叶子的正常形态发育是必需的。生长素响应因子(ARF)是一类调节植物生长、发育的转录因子家族, 先前的研究已经证明 ARF 中的一些成员为 miR160 的靶基因, 如 ARF10, ARF16, ARF17^[36], ARF6 和 ARF8 是 miR167 的靶基因^[37]。同样本研究发现, 小麦 miR160 也靶定 ARF10, 而 miR167 同样靶定 ARF8。Mallory 和 Barte^[33]通过在拟南芥中过量表达对 miR160 有抗性的 ARF17 来研究 miR160 在植物生长

发育过程中的作用, 结果转基因植物表现为积累 ARF17mRNA 含量和根生长的缺陷。这表明这一机制对植物根的正常发育起关键作用。

另外, 本研究预测到 miR172 靶定 APETALA2 及其同源的两个蛋白(TC209305, TC209306)。APETALA2 (APE2)家族属于调控花组织正常发育的 ABC 模型中 A 家族。过量表达对 miR172 抗性的 APE2 的转基因拟南芥会出现早花的现象^[38]。除了 APE2, miR172 还会靶定 APE2 的同源基因, 如拟南芥中 TARGET OF EAT1 (TOE1), TOE2, TOE3 和玉米中的 INDETERMINATE SPIKE-LET1 和 GLOY15^[39,40]。所以本研究预测的两个 APE2 的同源基因(TC209305, TC209306)也极有可能是小麦 miR172 的靶基因, 与 APE2 一起调控着小麦的正常开花时间。

本研究还预测到 NAC 转录因子家族中的两个成员(NAC1 和 NAC2)是小麦 miR164 的靶基因, 在野生拟南芥中由 35S 启动子引导的过量表达 miR164 会导致花组织的融合和子叶的融合^[41]。说明由 miR164 诱导的 NAC 转录因子的降解对植物花的正常发育是必需的。除了靶定转录因子外, 小麦 miRNA 的靶基因还编码一些参与新陈代谢、胁迫反应等的蛋白。如, 小麦 miR818 的靶基因编码运输蛋白 SEC61 的 β 亚体, 另外 miR818 还会靶定一些与疾病相关的蛋白。

参考文献

- 1 Bartel D P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 2004, 116: 281—297[DOI]
- 2 Jones-Rhoades M W, Bartel D P, Bartel B. MicroRNAs and their regulatory roles in plants. *Annu Rev Plant Biol*, 2006, 57: 19—53[DOI]
- 3 Papp I, Mette M F, Aufsatz W, et al. Evidence for nuclear processing plant microRNA and short interfering precursors. *Plant Physiol*, 2003, 132: 1382—1390[DOI]
- 4 Park M Y, Wu G, Gonzalez-Sulser A, et al. Nuclear processing and export of microRNAs in *Arabidopsis*. *Proc Natl Acad Sci USA*, 2005, 102: 3691—3696[DOI]
- 5 Arazi T, Talmor-Neiman M, Stav R, et al. Cloning and characterization of microRNAs from moss. *Plant J*, 2005, 43: 837—848[DOI]
- 6 Sunkar R, Zhu J K. Novel and stress-regulated microRNAs and other small RNAs from *Arabidopsis*. *Plant Cell*, 2004, 16: 2001—2019[DOI]
- 7 Sunkar R, Girke T, Jain P K, et al. Cloning and characterization of microRNAs from rice. *Plant Cell*, 2005, 17: 1397—1411[DOI]
- 8 Zhang B H, Pan X P, Cannon C H, et al. Conservation and divergence of plant microRNA genes. *Plant J*, 2006, 46: 243—259[DOI]
- 9 Lee R C, Feinbaum R L, Ambros V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, 1993, 75: 843—854[DOI]

- 10 Brown J R, Sanseau P. A computational view of microRNAs and their targets. *Drug Discov Today*, 2005, 10: 595—601 [\[DOI\]](#)
- 11 Lu S, Sun Y H, Shi R, et al. Novel and mechanical stress-responsive microRNAs in *Populus trichocarpa* that are absent from *Arabidopsis*. *Plant Cell*, 2005, 17: 2186—2203 [\[DOI\]](#)
- 12 Moxon S J, Szittyá R C, Schwach G, et al. Deep sequencing of tomato short RNAs identifies microRNAs targeting genes involved in fruit ripening. *Genome Res*, 2008, 18: 1602—1609 [\[DOI\]](#)
- 13 Qiu D Y, Pan X P, Wilson W I, et al. High throughput sequencing technology reveals that the taxoid elicitor methyl jasmonate regulates microRNA expression in Chinese yew (*Taxus chinensis*). *Gene*, 2009, 436: 37—44 [\[DOI\]](#)
- 14 Jones-Rhoades M W, Bartel D P. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol Cell*, 2004, 14: 787—799 [\[DOI\]](#)
- 15 Wang X J, Reyes J L, Chua N H, et al. Prediction and identification of *Arabidopsis thaliana* microRNAs and their targets. *Genome Biol*, 2004, 5: R65 [\[DOI\]](#)
- 16 Bonnet E, Wuyts J, Rouze P, et al. Detection of 91 potential conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* identifies important targets. *Proc Natl Acad Sci USA*, 2004, 101: 11511—11516 [\[DOI\]](#)
- 17 Zhang B H, Pan X P, Wang Q L, et al. Identification and characterization of novel plant microRNAs using EST analysis. *Cell Res*, 2005, 15: 336—360 [\[DOI\]](#)
- 18 Zhang B H, Pan X P, Anderson T A. Identification of 188 conserved maize microRNAs and their targets. *FEBS Lett*, 2006, 580: 3753—3762 [\[DOI\]](#)
- 19 Qiu C X, Xie F L, Zhu Y Y, et al. Computational identification of microRNAs and their targets in *Gossypium hirsutum* expressed sequence tags. *Gene*, 2007, 395: 49—61 [\[DOI\]](#)
- 20 Jin W, Li N, Zhang B, et al. Identification and verification of microRNA in wheat (*Triticum aestivum*). *J Plant Res*, 2008, 121: 351—355 [\[DOI\]](#)
- 21 Yao Y, Guo G, Ni Z, et al. Cloning and characterization of microRNAs from wheat (*Triticum aestivum* L.). *Genome Biol*, 2007, 8: R96 [\[DOI\]](#)
- 22 Dryanov A, Zakharov A, Gulick P J. Data mining for miRNAs and their targets in the Triticeae. *Genome*, 2008, 51: 433—443 [\[DOI\]](#)
- 23 Griffiths-Jones S, Grocock R J, van Dongen S, et al. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res*, 2006, 34: D140—D144 [\[DOI\]](#)
- 24 Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*, 2003, 31: 3406—3415 [\[DOI\]](#)
- 25 Meyers B C, Axtell M J, Bartel B, et al. Criteria for annotation of plant microRNAs. *Plant Cell*, 2008, 20: 3186—3190 [\[DOI\]](#)
- 26 Corpet F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res*, 1988, 16: 10881—10890 [\[DOI\]](#)
- 27 Zhang B H, Pan X P, Cox B, et al. Evidence that miRNAs are different from other RNAs. *Cell Mol Life Sci*, 2006, 63: 246—254 [\[DOI\]](#)
- 28 Yu J, Wang F, Yang G H, et al. Human microRNA clusters: genomic organization and expression profile in leukemia cell lines. *Biochem Biophys Res Commun*, 2006, 349: 59—68 [\[DOI\]](#)
- 29 Talmor-Neiman M, Stva R, Frank W, et al. Novel microRNAs and intermediates of microRNA biogenesis from moss. *Plant J*, 2006, 47: 25—37 [\[DOI\]](#)
- 30 Lu S, Sun Y H, Amerson H, et al. MicroRNAs in loblolly pine (*Pinus taeda* L.) and their association with fusiform rust gall development. *Plant J*, 2007, 51: 1077—1098 [\[DOI\]](#)
- 31 Allen E, Xie Z, Gustafson A M, et al. MicroRNA-directed phrasing during trans-acting siRNA biogenesis in plants. *Cell*, 2005, 121: 207—221 [\[DOI\]](#)
- 32 Guo H S, Xie Q, Fei J F, et al. MicroRNA directs mRNA cleavage of the transcription factor NAC1 to downregulate auxin signals for *Arabidopsis* lateral root development. *Plant Cell*, 2005, 17: 1376—1386 [\[DOI\]](#)

- 33 Mallory A C, Bartel D P. MicroRNA-directed regulation of *Arabidopsis* AUXIN RESPONSE FACTOR17 is essential for proper development and modulate expression of early auxin response genes. *Plant Cell*, 2005, 17: 1360—1375 [\[DOI\]](#)
- 34 Palatnik J F, Allen E, Wu X, et al. Control of leaf morphogenesis by microRNAs. *Nature*, 2003, 425: 257—263 [\[DOI\]](#)
- 35 Lee Y, Kim M, Han J, et al. MicroRNA genes are transcribed by RNA polymerase . *EMBO J*, 2004, 23: 4051—4060 [\[DOI\]](#)
- 36 Rhoades M W, Reinhart B J, Lim L P, et al. Prediction of plant microRNA targets. *Cell*, 2002, 110: 513—520 [\[DOI\]](#)
- 37 Bartel B, Bartel D P. MicroRNAs: at the root of plant development. *Plant Physiol*, 2003, 132: 709—717 [\[DOI\]](#)
- 38 Chen X. A microRNA as a translational repressor of APETALA2 in *Arabidopsis* flower development. *Science*, 2004, 303: 2022—2025 [\[DOI\]](#)
- 39 Aukerman M J, Sakai H. Regulation of flowering time and floral organ identity by a microRNA and its APETALA2-like target genes. *Plant Cell*, 2003, 15: 2730—2741 [\[DOI\]](#)
- 40 Park W, Li J, Song R, et al. CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. *Curr Bio*, 2002, 12: 1484—1495 [\[DOI\]](#)
- 41 Mallory A C, Dugas D V, Bartel D P, et al. MicroRNA regulation of NAC-domain targets is required for proper formation and separation of adjacent embryonic, vegetative, and floral organs. *Curr Biol*, 2004, 14: 1035—1046 [\[DOI\]](#)