Sequencing sliced ends reveals microRNA targets

Ian R Henderson & Steven E Jacobsen

Global sequencing of cleaved mRNAs enables identification of the targets of microRNA silencing.

Hundreds of microRNAs (miRNAs) have been identified in plants and animals by cloning and sequencing, but assigning these miRNAs to their cognate target genes remains technically challenging[1,2]. In this issue, German *et al.*[3] describe the use of high-throughput sequencing to globally sample degraded RNAs in *Arabidopsis thaliana* and thereby identify new miRNAs and their targets. Similar techniques have been developed in two other recent reports[4,5]; together, the three studies provide powerful new tools for miRNA discovery and the analysis of mRNA turnover.

Since their discovery nearly a decade ago[6], tiny RNA molecules have emerged as regulators of an astonishing number of processes. miRNAs, one prominent class of these RNAs, are endogenous ~21-nucleotide molecules that repress fully or partially complementary target mRNAs either through endonucleolytic cleavage (ARGONAUTE-mediated 'slicing') or through translational repression[1,2]. The assignment of miRNAs to their cognate target genes with computational approaches can be difficult when sequence complementarity is weak. An alternative approach in these cases is to search experimentally for the signatures of miRNA action.

The new studies take advantage of the fact that intact mRNA and cleaved fragments of mRNA have different structures at their 5′ ends (**Fig. 1a**). During transcription, mRNAs are modified with 7-methylguanosine via a 5′-5′ triphosphate linkage, a structure termed the 5′ cap. The 5′ cap protects mRNA from exonucleases and also promotes protein translation. In contrast, the cleavage fragment resulting from an endonucleolytic cut possesses a 5′ monophosphate rather than a 5′ cap (**Fig. 1a**). Conventional techniques for defining the 5′ ends of mRNA, such as those involving rapid amplification of cDNA ends (RACE), often use the cap as a molecular marker for the mRNA 5′ end. The new studies exploit the RACE principle but ignore the 5′ cap and selectively clone molecules with a 5′ monophosphate[3–5] (**Fig. 1a**).

The three approaches, called parallel analysis of RNA ends[3] (PARE), degradome sequencing[4] and genome-wide mapping of uncapped and cleaved transcripts[5] (GMUCT), sample the 3′ cleavage fragments produced by endonucleolytic cuts, capturing a global snapshot of degraded RNAs. The libraries of cut ends are then sequenced by high-throughput methods and the reads mapped back to the *Arabidopsis* genome[3–5]. Short sequence reads of ~20 base pairs are usually sufficient to map most reads to a specific mRNA transcript.

Analysis of these degradome libraries revealed that the majority of expressed genes were represented[3–5]. The degree to which a gene was sequenced was found to correlate with mRNA abundance[3–5]. This suggests that some level of endonucleolytic cleavage and turnover is the norm for most mRNAs. Further support of this idea came from degradome sequencing of an *xrn4/ein5* mutant[3,5]. XRN4/EIN5 is a 5′-3′ exonuclease that degrades the 3′ fragments of cleaved mRNAs[7,8]. Hence, an *xrn4/ein5* mutant should stabilize these cleavage

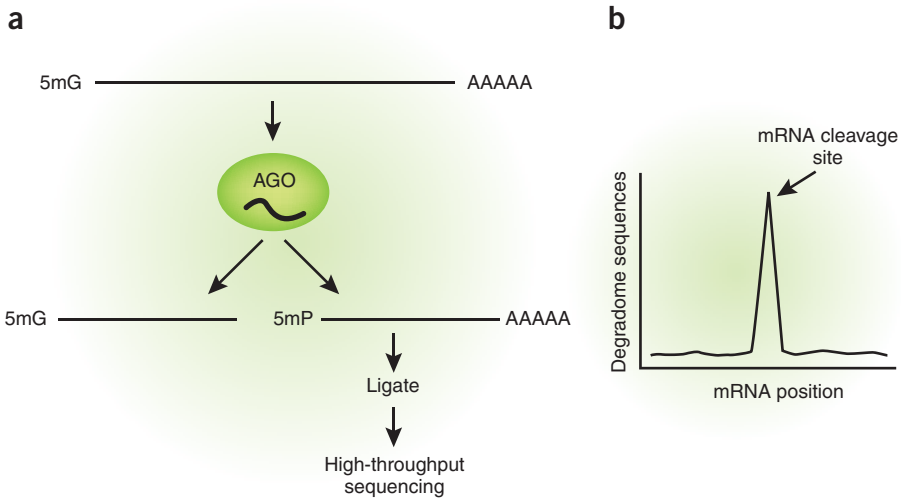

Figure 1 Discovery of miRNAs by sequencing cleaved mRNA fragments. (**a**) An intact mRNA possesses a 5′ cap (5mG, 5′ 7-methylguanosine) structure and a 3′-poly(A) tail. The presence of an miRNA (loaded into an ARGONAUTE protein) with complementarity to the mRNA can lead to endonucleolytic cleavage, yielding a downstream fragment with a 5′ monophosphate end. Molecules with a 5′ monophosphate are ligated and used for high-throughput sequencing. AGO, ARGONAUTE. (**b**) Endonucleolytic cleavage directed by the miRNA is evident as a peak in the sequenced mRNA fragments.

Ian R. Henderson and Steven E. Jacobsen are at the Department of Molecular, Cell and Developmental Biology, University of California, Los Angeles, Los Angeles, California 90095, USA. Steven E. Jacobsen is also a member of the Howard Hughes Medical Institute.
email: jacobsen@ucla.edu

fragments and make them easier to detect using sequencing. Indeed, this turned out to be the case, as many genes showed greater sequence representation in an *xrn4/ein5* background[3].

Sequencing of small RNAs in the *xrn4/ein5* mutant also showed that a number of new small RNA peaks appeared in genes that were not observed in wild type[5]. This is probably due to accumulation of uncapped fragments of mRNA, which can serve as substrates for an RNA-dependent RNA polymerase to generate double-stranded RNA[8]. This double-stranded RNA can then act as a substrate for the generation of siRNAs by the RNA interference pathway[8]. Together, the three studies illustrate how a combination of mutations in RNA metabolism and high-throughput sequencing technologies can be used to gain global insights into RNA dynamics.

German *et al.*[3] could use PARE to identify miRNA targets because miRNAs can cause endonucleolytic cleavage of mRNA, especially in cases of perfect complementarity[1,2]. This means that miRNA activity should be detectable as peaks of degradome signatures at the cleavage site (**Fig. 1b**). Indeed, the majority of known miRNA targets were easily identified by searching for such peaks[3–5]. Cloning and sequencing of the cut ends of mRNA also captures information about how frequently and accurately specific target sites are pinpointed for cleavage by miRNAs[3–5]. To discover unknown miRNAs, German *et al.*[3] searched for peaks not accounted for by known miRNAs. Several suggestive peaks were found to be accurate proxies for the presence of novel miRNAs[3]. This is surprising given that *Arabidopsis* miRNAs have already been studied in depth[2] and underscores the potential of these methods for miRNA discovery in other less-characterized species. In addition to identifying new miRNAs, German *et al.*[3] also showed that at least one miRNA precursor transcript is a target of its own miRNA. Finally, this technique allows identification of other small RNAs that cause mRNA cleavage, as highlighted by the degradome signatures generated by *trans*-acting siRNAs (tasiRNAs)[3,4].

PARE provides a comprehensive means of analyzing patterns of RNA degradation that has many advantages, not the least of which is the discovery of new miRNAs. Searching for the signatures of cleavage may be a useful tool to help identify miRNAs with low abundance or difficult annotation. Interestingly, translational suppression has recently been shown to be a common feature of *Arabidopsis* miRNA function, occurring simultaneously with mRNA cleavage and irrespective of target mRNA complementarity[9]. Understanding of the breadth of this translational miRNA pathway should benefit from the development of similarly comprehensive genomics approaches. Degradome sequencing also provides an attractive means of testing the relative importance of endonucleolytic cleavage versus translational suppression mechanisms for miRNAs in other organisms, for example, in animals where translational suppression is thought to dominate[1].

1. Bartel, D.P. *Cell* **116**, 281–287 (2004).
2. Chapman, E.J. & Carrington, J.C. *Nat. Rev. Genet.* **8**, 884–896 (2007).
3. German, M.A. *et al. Nat. Biotechnol.* **26**, 941–946 (2008).
4. Addo-Quaye, C., Eshoo, T.W. Bartel, D.P. & Axtell, M.J. *Curr. Biol.* **18**, 758–762 (2008).
5. Gregory, B.D. *et al. Dev. Cell* **14**, 854–866 (2008).
6. Hamilton, A.J. & Baulcombe, D.C. *Science* **286**, 950–952 (1999).
7. Souret, F.F. *et al. Mol. Cell* **15**, 173–183 (2004).
8. Gazzani, S. *et al. Science* **306**, 1046–1048 (2004).
9. Brodersen, P. *et al. Science* **320**, 1185–1190 (2008).

Nematology: terra incognita no more

James P McCarter

The genome sequence of the plant-parasitic roundworm *Meloidogyne incognita* opens new avenues to boosting food production.

If we could select a single agricultural pest to banish, the root-knot nematode would be an excellent choice. Few plant pathogens cause such substantial yield losses, invade so many crop hosts and are as difficult to control. The report of the assembled genome of the root-knot nematode species *M. incognita* by Abad *et al.*[1] in this issue provides new insights into this enemy and should aid the development of environmentally sustainable nematicides and transgenic crops that target several important agricultural pests. This study also represents the first genome sequence of a metazoan plant pathogen and of an animal that reproduces asexually. These two features appear to have striking effects on genome composition, including the presence of 61 cell wall–degrading enzymes—a number unprecedented in any animal—and of divergent ancient allelic regions that evolve without recombination.

Root-knot nematodes are microscopic worms that enter plant roots as larvae and migrate between cells toward a vascular feeding site. There, they induce formation of multinucleated plant cells that serve as a nutritional sink. Sedentary female worms remain at this site, molting to adults and laying eggs that hatch, repeating the infection cycle. Surrounding plant cells form visible knots, and plant growth and yield are impaired as nutrients are diverted to the worms. The stunted root systems have diminished water transport capacity, making the plant especially vulnerable to drought, which further reduces crop yields.

Although root-knot nematodes are responsible for ~5% of crop losses worldwide, technology for controlling them has advanced little in the last 30 years. Nematicides such as organophosphates are nonspecific neurotoxins that are as lethal to humans, birds and fish as they are to worms, and most have been withdrawn from the market or greatly restricted in use. The ozone-depleting fumigant methyl bromide is also being phased out. Thus, control of plant-parasitic nematodes, especially root-knot nematodes, represents a major opportunity for improving crop yields.

Nematodes are the most abundant animals on earth and an ancient phylum in which parasitism of plants and animals, including humans, has arisen independently at least seven times[2]. Therefore, it was perhaps only fitting that the first animal genome to be sequenced was that of the model nematode *Caenorhabditis elegans*[3]. In the ensuing decade, nematode genomics has continued to yield surprises. Gene collections from sampled species are remarkably diverse, with >4,000 nematode-specific protein families identified[4]. The human parasite *Brugia malayi*, a filarial roundworm responsible for elephantiasis, was shown to harbor, in

James P. McCarter is at Divergence Inc., 893 North Warson Road, St. Louis, Missouri 63141, USA and the Washington University Genome Center, 4444 Forest Park Blvd., St. Louis, Missouri 63108, USA.
e-mail: mccarter@divergence.com