related, the three small RNAs probably correspond to different stages in the evolution of functional miRNA genes (**Fig. 1**).

According to this new model (Fig. 1), an initial inverted duplication brings the miRNA founder gene and related members under negative regulation by short interfering RNAs covering the entire duplicated region (step 1). Limited sequence divergence progressively occurs under constraints that maintain the foldback structure, such that DCLs other than DCL-1 produce discrete small RNA species, possibly lacking trans-acting activity. This leads to 'transitional forms' akin to ASRP1729 (step 2). Further sequence divergence and adaptation to DCL-1 recognition and processing generates young miRNA genes, such as MIR161 and MIR163, producing a uniform small RNA population (step 3). These in turn evolve to a point that only the miRNA sequence resembles the founder gene sequence, as seen in most older miRNA genes (step 4).

Implications and uncertainties

This scenario provides a simple framework in which retention or loss of small RNA complementary sites among members of the founder gene family creates differential post-transcriptional regulations and thus new functions. The model is unlikely to apply in metazoans, however, where mRNAs often have several complementary sites to many distinct miRNAs within their 3' untranslated regions, allowing finetuned regulation¹² as opposed to the on-off control seen in plants. Thus, gain-of-function mutations affecting target 3' untranslated regions, rather than *de novo* formation of miRNA genes, probably account for new miRNA-directed regulatory pathways in animals. The difference might explain the absence of common miRNA or miRNA targets between plants and animals and suggests that plant miRNA genes may have arisen independently, after the split from a common lineage.

Several uncertainties remain. First, small RNA loci corresponding to step 1 have not been detected in plants. Second, the model relies heavily on the idea that miRNA gene evolution is guided by mere structural constraints, yet the influence of bulges and mismatches of plant miRNA precursors on DCL-1 activity is entirely unknown. Third, given the scope of the model, the actual number of inverted or duplicated small RNA loci seems to be relatively modest. Therefore, some miRNA loci might have evolved by different mechanisms. Alternatively, the small RNA database may have been strongly biased by sequences from highly abundant, silencing small RNAs found in *A. thaliana*, which are not produced from foldback transcripts but rather by the action of RNA-dependent RNA polymerases^{3,13,14}. Fourth, the study was carried out with RNA from healthy, normally grown plants. Given their highly evolving nature, young small RNAs are ideally suited to convey rapid adaptation to the environment and, therefore, will probably be more abundant in plants subjected to stress. In this respect, gene families rapidly evolving under stress constraints, such as complex disease resistance gene clusters¹⁵, are prime candidates for young small RNA-generating loci in plants.

- Carrington, J.C. & Ambros, V. Science 301, 336–338 (2003).
- 2. Baulcombe, D. Nature 431, 356-363 (2004).
- 3. Xie, Z. et al. PLoS Biol. 2, 642-652 (2004).
- 4. Bartel, D.P. Cell 116, 281–297 (2004).
- Rhoades, M.W. *et al. Cell* **110**, 513–520 (2002).
 Parizotto, E.A. *et al. Genes Dev.* **18**, 2237–23.
- 5. Parizotto, E.A. *et al. Genes Dev.* **18**, 2237–2242 (2004).
- Mansfield, J.H. *et al. Nat. Genet.* **36**, 1079–1083 (2004).
- 8. Allen, E. et al. Nat. Genet. 36,1282-1290 (2004).
- 9. Boutet, S. et al. Curr. Biol. 13, 843-848 (2003).
- 10. Llave, C. et al. Science 297, 2053–2056 (2002).
- 11. Kasschau, K.D. et al. Dev. Cell 4, 205-217 (2003).
- 12. Bartel, D.P. & Chen, C.Z. Nat. Rev. Genet. 5, 396–400 (2004).
- 13. Peragine, A. *et al. Genes Dev.* **18**, 2368–2379 (2004).
- 14. Vazquez, F. et al. Mol. Cell. 16, 69-79 (2004).
- 15. Leister, D. Trends Genet. 3, 116-122 (2004).

Exploring nematode diversity

Thomas Blumenthal & Richard E Davis

A massive study of expressed-sequence tags from both free-living and parasitic nematodes identifies tens of thousands of genes not present in *Caenorhabditis elegans*. Half of these are unique to nematodes and should provide new insights into the lifestyles of parasitic nematodes.

All nematodes are round worms and share a similar body plan, despite varying in length from ~100 μ m to more than 6 m (**Fig. 1**). Given this morphological similarity, one might think that all nematodes would express similar sets of proteins, but Parkinson *et al.*¹ report finding great diversity in expressed sequences among nematode species.

Genetic diversity in the nematode phylum was suggested previously by analysis of the genomic sequence of *Caenorhabditis briggsae*,

Thomas Blumenthal and Richard E. Davis are in the Departments of Biochemistry and Molecular Genetics and Pediatrics, University of Colorado School of Medicine, Aurora, Colorado 80045, USA.

e-mail: Tom.Blumenthal@UCHSC.edu

estimated to have diverged from *C. elegans* ~100 million years ago^2 . Although these two species are widely regarded as sister species (they are morphologically indistinguishable to the naive observer), they each have at least 2,000 genes (~10%) that are not found in the other species.

Now Parkinson *et al.*¹ report the sequencing of a large collection of expressed-sequence tags (>250,000) from 30 different nematode species across the phylum, including four of the five major clades. This report represents the most complete analysis of the genetic diversity of a single phylum yet attempted, and it is full of surprises. The expressedsequence tags defined ~93,000 genes, grouped into ~60,000 gene families. This shows that the *C. elegans* genome, which has only ~20,000 genes in 12,000 gene families, contains only a relatively small fraction of the genetic diversity of nematodes. Because only 30 of the estimated 100,000–1,000,000 species of nematodes on earth were sampled in this study, continued analysis of the phylum will probably uncover even greater diversity.

Thousands of new genes

Of the 93,000 genes found, more than half were unique to nematodes, with no known homologs. Only 15% were found in all four sampled clades; as expected, most of these are found outside the phylum as well. These conserved genes are mostly responsible for basic cellular processes. There was also a strong correlation between genes that are conserved throughout the phylum and those that result in an abnormal phenotype when subjected to RNA interference, providing further evidence of their fundamental importance. More than 1,200 genes found throughout the phylum were not found elsewhere. The authors identify these nematode-specific genes as potential pan-nematode drug and vaccine targets, because they may be important in conferring functions specific to this phylum.

Even though 30 different species were sampled, 30-50% of the sequences examined were found to be species-specific. Nearly half of the nematode-specific genes were found in more than one lineage. New genes appear at every node in the nematode tree, with some appearing very early in nematode evolution and many others appearing after the most recent branching.

These data point to the need for sampling a much wider variety of species than the few model organisms selected by investigators for intense study. It is misleading to base conclusions about protein and proteome diversity on complete sequences from a few very diverse species, when even members of the same phylum turn out to be so different from one another. Evolution is a continuous and continuing process, and new proteins are appearing rapidly in nematodes.

Parasitic lifestyles

Nematodes exploit a wide range of ecological niches, as both free-living and parasitic species. Parasitic nematodes infect 3 billion people worldwide, leading to considerable morbidity; they are a problem for livestock and domestic animals; and they result in billions of dollars in annual crop damage. Of the 30 different nematode species sampled in the study, 28 are animal or plant parasites. It is possible that the diversity of new genes sampled is at least partly a result of independent selection for parasitic lifestyle.

Parkinson et al.¹ show that although C. elegans has provided a wealth of molecular information about nematodes, there is much more to be learned by studying additional nematode species. This study indicates that parasitic nematodes of agricultural, veterinary and medical importance have unique genes that are not present in free-living C. elegans, the major nematode model.



Figure 1 Diversity of nematode morphology and lifestyle. (a) Adult Ascaris worm, an intestinal parasite of vertebrates. The human form is ~30 cm long and 0.5 cm wide. (b) Culture of Caenorhabditis, a freeliving nematode, illustrating life-cycle stages. Adults are 1.5 mm long and 80 µm wide. Photo by A. L. Diogo (University of Evora, Evora, Portugal). (c) Scanning electron micrograph of a female root-knot nematode (plant parasite), Meloidogyne arenaria. Photo by J. D. Eisenback (Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA) (d) Anterior end of hookworm Ancylostoma. Photo by J. D. Eisenback (Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA) (e) Posterior end (bursa) of hookworm Ancylostoma. (f) Scanning electron micrograph of anterior of Acrobeles sp., a bacterial feeder. Photo by P. Vieira (University of Evora, Evora, Portugal). (g) Trichinella larvae encysted in vertebrate muscle cells. Photo by J. D. Eisenback (Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA) (h) Microfilaria (larval stage) of filarial worm Wuchereria in blood smear. (i) Plant parasite Pratylenchus feeding inside cortical cell of root. Photo by U. Zunke (University of Hamburg, Hamburg, Germany).

Anecdotal evidence from laboratories studying parasitic nematodes suggests that their protein function, physiology and pharmacology may be quite different from those of C. elegans, implying a need for development of new molecular genetic and other experimental tools in parasitic nematodes. Preliminary studies describing the use of RNA interference and nucleic acid transfection in parasitic nematodes have recently been reported^{3,4}, but these adaptations of standard techniques need further refinement. Nematodes that offer unique biochemical systems and developmental advantages, such as Ascaris, can be exploited for experiments

that cannot be done in C. elegans or other parasitic species^{5–7}. Finally, the high level of nematode gene diversity argues for sequencing of the genomes of additional parasitic nematodes to explore the breadth of genes and search for new drug or vaccine targets.

- 1. Parkinson, J. et al. Nat. Genet. 36, 1259-1267 (2004)
- 2. Stein, L. et al. PLoS Biol. 1, 166-192 (2003).
- 3. Aboobaker, A.A. & Blaxter M.L. Trends Parasitol. 20, 178-184 (2004).
- Δ Boyle, J.P. & Yoshino, T.P. Int. J. Parasitol. 33, 1259-1268 (2003).
- Hannon, G.J. et al. Cell 61, 1247-1255 (1990). Maroney, P.A. et al. Proc. Natl. Acad. Sci. USA 87, 6.
- 709-713 (1990). 7. Maroney, P.A. et al. RNA 1, 714-723 (1995).