Caenorhabditis elegans functional genomics: Omic resonance

Jonathan Astin, Alyce Merry, Jeena Rajan and Patricia E. Kuwabara

Date received (in revised form): 20th February 2004

Jonathan Astin
is a PhD student in the Department of Biochemistry at the University of Bristol, UK.

Alyce Merry
is a PhD student in the Department of Biochemistry at the University of Bristol.

Jeena Rajan
is a computer biologist at the Wellcome Trust Sanger Institute in Cambridge, UK.

Patricia E. Kuwabara
is the William P. Coldrick Professor of Genomics in the Department of Biochemistry, University of Bristol. She is an MRC senior fellow.

Keywords: Caenorhabditis elegans, Caenorhabditis briggsae, *functional genomics, genome, proteomics, RNA*

Abstract

The nematode *Caenorhabditis elegans* is widely used as a model organism for studying many fundamental aspects of development and cell biology, including processes underlying human disease. The genome of *C. elegans* encodes over 19,000 protein-coding genes and hundreds of non-coding RNAs. The availability of whole genome sequence has facilitated the development of high throughput techniques for elucidating the function of individual genes and gene products. Furthermore, attempts can now be made to integrate these substantial functional genomics data collections and to understand at a global level how the flow of genomic information that is at the core of the central dogma leads to the development of a multicellular organism.

INTRODUCTION

The nematode *Caenorhabditis elegans* has been referred to as 'a worm' for all reasons. Sydney Brenner was the first to recognise these reasons and to select *C. elegans* as a model organism for biomedical studies, in 1965.[1] The *C. elegans* community has since grown to include over 500 laboratories. What makes this tiny worm so attractive? From the laboratory standpoint, it is easily and inexpensively propagated, readily stored and resuscitated from liquid nitrogen and suited for high throughput genomic analyses. At the experimental level, many researchers have been lured by the powerful genetics of *C. elegans*, the availability of a complete cell lineage of 959 somatic cells, a neuronal wiring diagram and the intrinsic advantages associated with studying organogenesis in a simple multicellular organism.[2,3] Research using the worm has made it possible to map genetic pathways and networks and to dissect their underlying molecular mechanisms. In turn, investigators have gained fundamental insights into basic cell biological processes such as ageing, apoptosis, behaviour, cell-to-cell signalling, cell division, embryogenesis, neuromuscular development, neuronal guidance and sex determination, to list but a few.[4] The *C. elegans* researchers Sydney Brenner, John Sulston and H. Robert Horvitz were jointly awarded the Nobel Prize for Physiology or Medicine in 2002, in recognition of their discoveries about cell lineage and programmed cell death.

The worm continues to thrive. In 1998, *C. elegans* gained the distinction of being the first multicellular organism with a fully sequenced genome.[5] Moreover, the *C. elegans* sequence is truly complete — covering each chromosome from end to end (Alan Coulson, personal communication). The present goal of researchers is to combine the experimental strengths of *C. elegans* with the availability of the genome sequence — and all of its associated -omic disciplines — in order to achieve an integrated understanding of organismal physiology, behaviour and development. In this paper, an overview of the progress that has been achieved to date is presented.

Patricia E. Kuwabara,
Department of Biochemistry,
University of Bristol,
The School of Medical Sciences,
University Walk,
Bristol BS8 1TD, UK

Tel: +44 (0) 117 331 7275
Fax: +44 (0) 117 928 8274
E-mail: p.kuwabara@bristol.ac.uk

THE GENOME AND GENE PREDICTIONS

The availability of the complete genome sequence of *C. elegans* is leading to an integrated understanding of whole organism physiology

The 100 megabase pair (Mbp) genome of *C. elegans* is currently predicted to encode ~19,000 protein-coding genes.[6] Gene annotations and information about many other aspects of *C. elegans* can be obtained on the *C. elegans* database, WormBase, and at a website maintained by Leon Avery.[7] Over one-half of all *C. elegans* genes have been experimentally verified from a very large set of expressed sequence tag (EST) sequences generated primarily by the group of Yuji Kohara.[8] Marc Vidal's group has confirmed an additional 4,000 genes by individually cloning predicted full-length protein-coding open-reading frames (ORFs) as part of a project coined 'the ORFeome'.[9–11]

The *C. elegans* genome also carries genes for an ever-increasing number of non-coding RNA (ncRNA) products, such as transfer RNA (tRNA), ribosomal RNA (rRNA), small nucleolar RNA (snoRNA) and small cytoplasmic RNA (scRNA). More recently, large families of non-coding microRNAs (miRNAs), tiny non-coding RNAs (tncRNAs) and endogenous small interfering RNAs (siRNAs), which might participate in endogenous RNA-mediated interference (RNAi), have been described. Many of the more than 100 predicted miRNAs are phylogenetically conserved;[12] however, the founding members of this family, *lin-4* and *let-7*, were first detected in *C. elegans*.[13,14] The miRNAs and tncRNAs are likely to play regulatory roles through non-RNAi mechanisms.[15,16]

THE PHENOME
RNA-mediated interference

The discovery of RNAi and its ease of application in *C. elegans* is facilitating whole genome annotation

One major goal of functional genomics is to understand the function of every gene in the genome. This is clearly a formidable task, but it is greatly aided by the availability of RNAi, which is a gene-silencing phenomenon related to post-transcriptional gene silencing in plants and quelling in *Neurospora*.[17] Introduction of gene-specific double-stranded RNA (dsRNA) causes degradation of its corresponding cognate mRNA and leads to a null or near null mutant phenocopy. RNAi is hypothesised to be part of a natural surveillance system that protects the worm against viruses and transposons — *C. elegans* researchers are the pirates who hijacked this natural defence mechanism for the cause of functional genomics.[18,19]

In *C. elegans*, dsRNA has traditionally been delivered by injection into the syncytial gonad;[19] however, RNAi can also be elicited simply by soaking worms in dsRNA solutions or by feeding them bacteria expressing dsRNA.[20,21] The ease with which dsRNA can be delivered, coupled with the complete genome sequence, makes *C. elegans* particularly amenable for performing systematic large-scale genomic RNAi screens.

Four datasets featuring genome-wide targeted disruption of *C. elegans* genes were presented in 2000. Gonczy *et al.*[22] systematically disrupted 2,174 genes, representing 96 per cent of the predicted ORFs on chromosome III by microinjecting dsRNA and using time-lapse differential interference contrast (DIC) microscopy to identify the genes involved in early embryonic cell divisions. Piano *et al.*[23] also screened for genes involved in embryogenesis, by injecting dsRNA prepared from complementary DNAs (cDNAs) enriched in the ovary. Maeda *et al.*[24] selected 2,479 cDNA clones that were part of the Kohara collection and scored the RNAi phenotypes of worms soaked in dsRNA solutions prepared from these clones. Fraser *et al.*[25,26] cloned 2,416 exon-rich gene polymerase chain reaction (PCR)-generated fragments into a vector capable of expressing dsRNA and produced an RNAi feeding library representing 87 per cent of the genes on chromosome I. Together, these four studies analysed almost one-third of the predicted genes in the *C. elegans* genome and discovered phenotypes for more than 888 genes, most of which were previously unknown. It is further expected that these results should aid in the annotation of other

genomes, because 60 per cent of the worm genes showing RNAi mutant phenocopies have homologues in other organisms.

New members of genetic pathways and networks are being revealed through directed genetic screens using RNAi

Despite the efficiency of detecting mutant phenocopies by microinjection with RNAi, it is losing favour over the soaking and feeding methods, which are less technically demanding and time-consuming. The feeding bacterial library, in particular, is a renewable resource, and an RNAi feeding library representing 87 per cent of the predicted *C. elegans* genes is available to researchers from the Ahringer group.[27,28] Using this library, Kamath *et al*.[29] screened worms for sterility, embryonic or larval lethality, delayed post-embryonic growth or other post-embryonic defects. Phenotypes were observed for around 10 per cent of genes, of which two-thirds had previously had no assigned function. The RNAi library has subsequently been used by a number of other researchers in more directed screens. Ashrafi *et al*.[30] performed a whole-genome RNAi screen to detect worms with disrupted fat storage. A total of 417 fat storage genes were identified: 150 of these have potential mammalian homologues that could be implicated in human obesity and associated disorders.[30] Murphy *et al*.[31] combined DNA microarray analysis with RNAi to screen for genes capable of prolonging animal longevity when inhibited. Many of the genes having this property were found to function in mitochondria, lending support to the hypothesis that there is a link between longevity and metabolism.[31,32]

A complete description of gene function is dependent on obtaining gene knockouts

There are caveats to using RNAi. RNAi does not always mimic the null phenotype, and some tissues appear to be resistant to RNAi — particularly in the nervous system.[33] These factors contribute to the generation of false-negative results. To counter these problems, RNAi can be performed in strains that are hypersensitive to RNAi, such as *rrf-3*, which can increase the ability to detect RNAi mutant phenotypes by 23 per cent.[34] The amount of inter-experimental variability in performing RNAi has been measured to range between 10 to 30 per cent; this observation highlights the importance of replication in RNAi studies. These considerations aside, RNAi has hugely enhanced the appeal of *C. elegans* as a model organism for functional genomics.

Gene knockouts

To gain a complete description of gene function, the majority of researchers concur that it is necessary to obtain genetic mutants. The *C. elegans* Genetics Centre maintains and distributes stocks and is the central repository of more than 1,000 mutants.[35] For those attempting to generate mutants using classical screens, a large number of verified small nucleotide polymorphisms (SNPs) are reported on WormBase to facilitate gene mapping.[36] Many now prefer to obtain mutants using high throughput PCR to identify worms that carry transposon insertions or small gene-specific deletions generated by chemical mutagenesis.[37–40] A newly funded European Union-funded initiative will soon be launched to identify mutations caused by mobilisation of the *Drosophila mos-1* transposon in *C. elegans* (N. Tavernarakis, personal communication).[41]

THE TRANSCRIPTOME

DNA microarrays

Stuart Kim is responsible for pioneering the development of DNA microarrays for *C. elegans* gene expression profiling. The arrays developed by Kim and colleagues are composed of ~1 kilobase (kb) exon-rich PCR fragments, representing ~94 per cent of predicted *C. elegans* genes. The first DNA microarray study published by the Kim laboratory identified 875 genes that are enriched in the *C. elegans* germline.[42] The *C. elegans* female is actually a self-fertilising hermaphrodite that produces first sperm, then oocytes. The availability of genetic mutants made it possible to compare the expression profiles of mRNAs obtained from matched worm populations that lacked a germline or that produced only

sperm or only oocytes.[42] These comparisons led to the identification of 258 oocyte-enriched genes, 650 sperm-enriched genes and 508 genes that are intrinsic to the germline but are not sex specific.

DNA microarrays provide clues about the function of novel genes through guilt-by-association

One of the strengths of performing DNA microarray studies in *C. elegans* is the ability to validate target genes rapidly by RNAi. Two independent groups performed follow-up studies based on the data obtained above. Piano *et al.*[23] examined RNAi-generated mutant phenocopies produced by the oocyte-specific and germline-intrinsic genes with an interest in identifying genes involved in early embryogenesis. This study led to the development of a phenotypic classification system, which allows genes producing similar phenotypes to be clustered.[43] Colaiacovo *et al.*[44] sought to identify genes involved in meiosis; they performed RNAi using a subset of the germline genes identified by microarrays, which shared expression profiles with other genes already known to be involved in meiosis. Both Piano *et al.*[43] and Colaiacovo *et al.*[44] were successful in finding their genes of interest, and in showing that the germline-enriched genes identified from microarray analysis produced strong germline RNAi mutant phenocopies. cDNA subtraction followed by RNAi has also been used as a method to enrich for germline-enriched genes.[45]

RNAi provides a rapid method for validating target genes identified by DNA microarrays

Clever strategies have been devised to circumvent problems associated with isolating mRNA from tissues or cells that are limited in quantity or availability. By culturing *C. elegans*-dissociated embryos, it was shown that mRNA could be isolated and linearly amplified from a rare population of green fluorescent protein (GFP)-tagged neurones, corresponding to six of 3,000 adult nuclei, using fluorescence-activated cell sorting.[46] In a separate study, epitope-tagged poly(A)-binding protein, expressed under the control of a muscle-specific promoter, was used to co-precipitate mRNAs preferentially enriched in muscle.[47]

Microarrays have also been used as tools to gain clues about chromosomal organisation and *cis*-acting regulatory elements in *C. elegans*.[46–51] In addition, Kim *et al.*[48] showed that by pooling a large number of microarray datasets it is possible to generate a global three-dimensional expression map to visualise clusters of genes that appear to be co-regulated and to gain clues about the function of novel genes through guilt-by-association.

Hill *et al.*[52] were the first group to use a commercially-generated oligonucleotide-based array in order to measure temporal changes in transcriptional patterns. Such arrays make it possible to quantify mRNA levels. A gridded array carrying 23,166 oligonucleotides, which are nominally 65 nucleotides in length, will soon be available as a resource for the *C. elegans* community through the Genome Sequencing Center in St. Louis (E. Maris, personal communication).[53]

SAGE

Serial analysis of gene expression (SAGE) is a method that complements gene expression profiling with microarrays.[54] SAGE involves the generation, concatenation and sequencing of short — 14–21 base pair — tags uniquely corresponding to expressed sequences.[55] The number of times a specific SAGE tag is sequenced is directly proportional to the abundance of that tag within a population; thus, statistical significance is dependent on the depth of tag coverage. SAGE analysis has been performed primarily by researchers with access to high-capacity, low-cost DNA sequencing. The Genome Sequencing Center in British Columbia (BCGSC) used SAGE to compare the expression profiles of wild-type and long-lived developmentally arrested dauer worms.[56] The authors of this study were able to identify changes not previously detected using DNA microarrays, such as an abundantly transcribed telomerase-like

sequence and numerous anti-sense mitochondrial transcripts, which might be involved in endogenous RNAi.

Transcriptional reporters

Collections of GFP reporters are being generated to provide information about spatial and temporal patterns of gene expression

Several other methods are being used to develop an understanding of the temporal and spatial patterns of gene expression. Global projects are underway to develop GFP reporters driven by the predicted promoter region of individual genes, which in *C. elegans* is usually contained within a 3–5 kb region upstream of the initiation codon, as demonstrated by the Hope group.[57] It is anticipated that these studies will also lead to the identification of *cis*-acting regulatory elements. The BCGSC has already generated transgenic lines corresponding to over 1,000 *C. elegans* genes that have human homologues.[58] Other groups are using recombinational cloning coupled with biolistic transformation in an effort to improve the efficiency of generating transgenics (Ian Hope, personal communication).[59,60] Yuji Kohara's group has taken advantage of their large collection of ESTs and has used them as probes for *in situ* hybridisation.

THE PROTEOME

Yeast 2-hybrid interaction mapping is helping to place proteins in functional networks and to uncover interactions occurring between members of different networks

The aim of protein interaction mapping is to establish a network of physical interactions between proteins. Such a network can help to identify proteins participating in a common signalling pathway, reveal links between different pathways, provide information about macromolecular complex formation and help to assign function through guilt-by-association to proteins not yet annotated.

The yeast two-hybrid system is the most familiar method of interaction mapping in *C. elegans*, although the development of protein chips is on the horizon. The two-hybrid system is engineered such that successful transcriptional activation of one or more yeast reporters occurs only when there is a productive interaction between the bait — a protein fused to a transcriptional activation domain (AD) — and prey — a second protein fused to a DNA-binding domain (DB). The feasibility of performing protein interaction mapping at the level of the whole genome was tested by first generating an interaction matrix with 29 *C. elegans* genes known to be involved in vulval development, many of which are components of a Ras/Map kinase signalling cascade.[61] To facilitate the transfer of cDNA sequences from vector to vector, clones were generated using a phage recombination system that is commercially available as Gateway (Invitrogen).[62] Two novel interactions were detected, and the biological relevance of one was verified. The matrix only uncovered six out of 11 known interactions — a 50 per cent detection rate, however. False-negatives can arise when proteins, such as a membrane protein, fail to localise to the nucleus or when protein–protein contact is inhibited by steric hindrance of the fusion partners or by the absence of post-translational modification.

A total of 27 vulval genes were subsequently used as baits to screen a global cDNA library; 148 interactions were detected.[61] To navigate through this wealth of data and to assign a significance judgment to individual interactions, the authors proposed that interactions conserved in other organisms, termed 'interologs', should have a high significance likelihood, and that closed networks of interacting proteins might be indicative of macromolecular complexes or signalling pathways.[61] False-positives are a major problem in yeast two-hybrid library screens and can arise when the bait is capable of self-activating transcription or when interacting proteins are not normally found in the same developmental stage, tissue or intracellular compartment. An attempt was made to estimate the rate of false-positives by analysing the sets of interactions occurring between proteins associated with subunits of the *C. elegans* 26S proteasome.[63] These authors found that at least 80 per cent of the observed interactions between the subunits were compatible with the solved

crystal structure of the proteasome, giving hope that the majority of observed interactions detected by cDNA screens are biologically relevant.

Multiple projects have been initiated to study the *C. elegans* proteome

Boulton and colleagues addressed the problem of validating predicted protein interactions using RNAi.[64] Again, a matrix of 78 *C. elegans* genes, composed of three confirmed and 75 potential orthologues of genes known to participate in the DNA damage response (DDR) from other organisms, was used to identify potential pairwise interactions. The matrix screen identified 17 interologs, and a screen using 67 of the DDR genes as baits identified 125 novel interactions and produced two large interolog clusters of known checkpoint proteins and DNA repair proteins. RNAi revealed that 12 of the potential orthologues and 11 of the novel genes displayed mutant phenotypes associated with DDR deficiencies after γ-irradiation. Presently, the Vidal group is taking a complex systems approach to integrate multiple, independently collected datasets to increase the predictive value and significance of whole genome studies.[65]

In addition to two-hybrid interaction mapping, a number of other projects are underway to characterise the proteome. WormPD is a database providing useful information about *C. elegans* proteins; however, it is only available through Incyte for a fee. The Southeast Collaboratory of Structural Genomics (SECSG) is attempting to develop high throughput methods for expressing and crystallising proteins, with the goal of obtaining three-dimensional structures for *C. elegans* proteins.[66] Multidimensional liquid chromatography, coupled with tandem mass spectroscopy, has been shown to be an effective method for proteomic profiling of *C. elegans* protein mixtures.[67] The ability to quantify protein expression through metabolic labelling has also been successfully applied to *C. elegans*.[68] Moreover, the utility of developing GFP or red fluorescence protein-tagged antibodies through the phage-display system has been demonstrated for *C. elegans*.[69]

The genome of *C. briggsae*

The genome sequence of *C. briggsae* has improved the annotation of the *C. elegans* sequence through comparative genomics

The value of a genome sequence can be further enhanced through comparative genomics. A high-quality draft sequence of the genome of *C. briggsae*, a worm that is estimated to have diverged from *C. elegans* about 100 million years ago, makes it possible to improve genome annotations and to examine the nature of evolutionary change. *C. elegans* and *C. briggsae* are almost indistinguishable at the morphological level and they also share many similarities at the genomic level. A global analysis comparing and contrasting the two genomes has recently been published,[70] although a number of earlier studies have addressed questions related to the evolution of genomic information and gene function using 12 Mbp of finished *C. briggsae* sequence.[71–74] The paper by Stein *et al.*[70] indicates that Sydney Brenner selected the worm with the smaller genome; the 104 Mbp genome of *C. briggsae* is larger than the 100.3 Mbp of *C. elegans* because of an increase in non-conserved repetitive DNA. The estimated number of protein-coding genes is similar in both species, at about 19,000; however, 800 *C. briggsae* genes have no detectable matches in *C. elegans*. Nonetheless, most operons and non-coding RNAs have been retained by both genomes. In *C. elegans*, operons are composed of closely clustered genes that are transcribed into a polycistronic pre-mRNA.[49] Resolution of these pre-mRNAs into individual mRNAs is achieved through a process known as *trans*-splicing.[49] The challenge remains for researchers to extract information about *cis*-acting regulatory elements and to determine the extent to which gene function has been conserved between the two species. As with *C. elegans*, however, RNAi should prove to be an invaluable tool.

SUMMARY

The simplicity of *C. elegans* has been exploited to great advantage in an effort

to understand how its genomic blueprint controls organismal development. The field of *C. elegans* research remains vibrant and it is anticipated that the worm will lead to the holy grail of genomics — the complete structural and functional description of all genes.

References

1. Wood, W. B. and the Community of *C. elegans* Researchers (1988), 'The nematode *Caenorhabditis elegans*', Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, NY.
2. Sulston, J. E., Schierenberg, E., White, J. G. *et al.* (1983), 'The embryonic cell lineage of the nematode *Caenorhabditis elegans*', *Dev. Biol.*, Vol. 100, pp. 64–119.
3. White, J. G., Southgate, E., Thomson, J. N. *et al.* (1986), 'The structure of the nervous system of *Caenorhabditis elegans*', *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, Vol. 314, pp. 1–340.
4. Riddle, D. L., Blumenthal, T., Meyer, B. J. *et al.* (1997), '*C. elegans* II', Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, NY.
5. The *C. elegans* Sequencing Consortium (1998), 'Genome sequence of the nematode *C. elegans*: A platform for investigating biology', *Science*, Vol. 282, pp. 2012–2018.
6. URL: http://www.wormbase.org.
7. URL: http://elegans.swmed.edu/.
8. URL: http://nematode.lab.nig.ac.jp/.
9. Reboul, J., Vaglio, P., Tzellas, N. *et al.* (2001), 'Open-reading-frame sequence tags (OSTs) support the existence of at least 17,300 genes in *C. elegans*', *Nat. Genet.*, Vol. 27, pp. 332–336.
10. Reboul, J., Vaglio, P., Rual, J. F. *et al.* (2003), '*C. elegans* ORFeome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression', *Nat. Genet.*, Vol. 34, pp. 35–41.
11. URL: http://worfdb.dfci.harvard.edu/
12. Pasquinelli, A. E., Reinhart, B. J., Slack, F. *et al.* (2000), 'Conservation of the sequence and temporal expression of *let-7* heterochronic regulatory RNA', *Nature*, Vol. 408, pp. 86–89.
13. Lee, R. C., Feinbaum, R. L. and Ambros, V. (1993), 'The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*', *Cell,* Vol. 75, pp. 843–854.
14. Reinhart, B. J., Slack, F. J., Basson, M. *et al.* (2000), 'The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*', *Nature*, Vol. 403, pp. 901–906.
15. Ambros, V., Lee, R. C., Lavanway, A. *et al.* (2003), 'MicroRNAs and other tiny endogenous RNAs in *C. elegans*', *Curr. Biol.*, Vol. 13, pp. 807–818.
16. Grad, Y., Aach, J., Hayes, G. D. *et al.* (2003), 'Computational and experimental identification of *C. elegans* microRNAs', *Mol. Cell*, Vol. 11, pp. 1253–1263.
17. Bosher, J. M. and Labouesse, M. (2000), 'RNA interference: Genetic wand and genetic watchdog', *Nat. Cell Biol.*, Vol. 2, pp. E31–36.
18. Plasterk, R. H. (2002), 'RNA silencing: The genome's immune system', *Science,* Vol. 296, pp. 1263–1265.
19. Fire, A., Xu, S., Montgomery, M. K. *et al.* (1998), 'Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*', *Nature*, Vol. 391, pp. 806–811.
20. Tabara, H., Grishok, A. and Mello, C. C. (1998), 'RNAi in *C. elegans*: Soaking in the genome sequence', *Science*, Vol. 282, pp. 430–431.
21. Timmons, A. and Fire, A. (1998), 'Specific interference by ingested dsRNA', *Nature*, Vol. 395, p. 854.
22. Gonczy, P., Echeverri, C., Oegema, K. *et al.* (2000), 'Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III', *Nature*, Vol. 408, pp. 331–336.
23. Piano, F., Schetter, A. J., Mangone, M. *et al.* (2000), 'RNAi analysis of genes expressed in the ovary of *Caenorhabditis elegans*', *Curr. Biol.*, Vol. 10, pp. 1619–1622.
24. Maeda, I., Kohara, Y., Yamamoto, M. *et al.* (2001), 'Large-scale analysis of gene function in *Caenorhabditis elegans* by high-throughput RNAi', *Curr. Biol.*, Vol. 11, pp. 171–176.
25. Fraser, A. G., Kamath, R. S., Zipperlen, P. *et al.* (2000), 'Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference', *Nature,* Vol. 408, pp. 325–330.
26. Timmons, L., Court, D. L. and Fire, A. (2001), 'Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*', *Gene*, Vol. 263, pp. 103–112.
27. Kamath, R. S. and Ahringer, J. (2003), 'Genome-wide RNAi screening in *Caenorhabditis elegans*', *Methods*, Vol. 30, pp. 313–321.
28. URL: http://www.hgmp.mrc.ac.uk.
29. Kamath, R. S., Fraser, A. G., Dong, Y. *et al.* (2003), 'Systematic functional analysis of the

Caenorhabditis elegans genome using RNAi', *Nature*, Vol. 421, pp. 231–237.

30. Ashrafi, K., Chang, F. Y., Watts, J. L. *et al.* (2003), 'Genome-wide RNAi analysis of *Caenorhabditis elegans* fat regulatory genes', *Nature,* Vol. 421, pp. 268–272.
31. Murphy, C. T., McCarroll, S. A., Bargmann, C. I. *et al.* (2003), 'Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans*', *Nature*, Vol. 424, pp. 277–283.
32. Lee, S. S., Lee, R. Y., Fraser, A. G. *et al.* (2003), 'A systematic RNAi screen identifies a critical role for mitochondria in *C. elegans* longevity', *Nat. Genet.*, Vol. 33, pp. 40–48.
33. Tavernarakis, N., Wang, S. L., Dorovkov, M. *et al.* (2000), 'Heritable and inducible genetic interference by double-stranded RNA encoded by transgenes', *Nat. Genet.,* Vol. 24, pp. 180–183.
34. Simmer, F., Moorman, C., van der Linden, A. M. *et al.* (2003), 'Genome-wide RNAi of *C. elegans* using the hypersensitive *rrf-3* strain reveals novel gene functions', *PLoS Biol.*, Vol. 1, pp. 77–84.
35. URL: http://biosci.umn.edu/CGC/CGChomepage.htm
36. Wicks, S. R., Yeh, R. T., Gish, W. R. *et al.* (2001), 'Rapid gene mapping in *Caenorhabditis elegans* using a high density polymorphism map', *Nat. Genet.*, Vol. 28, pp. 160–164.
37. Zwaal, R. R., Broeks, A., van Meurs, J. *et al.* (1993), 'Target-selected gene inactivation in *Caenorhabditis elegans*, using a frozen transposon insertion mutant bank', *Proc. Natl. Acad. Sci. USA*, Vol. 90, pp. 7431–7435.
38. Jansen, G., Hazendonk, F., Tijjssen, K. L. and Plasterk, R. H. A. (1997), 'Reverse genetics by chemical mutagenesis in *Caenorhabditis elegans*', *Nat. Genet.*, Vol. 17, pp. 119–121.
39. Edgley, M., D'Souza, A., Moulder, G. *et al.* (2002), 'Improved detection of small deletions in complex pools of DNA', *Nucleic Acids Res.*, Vol. 30, p. e52.
40. URL: http://celeganskoconsortium.omrf.org/
41. Bessereau, J. L., Wright, A., Williams, D. C. *et al.* (2001), 'Mobilization of a *Drosophila* transposon in the *Caenorhabditis elegans* germ line', *Nature,* Vol. 413, pp. 70–74.
42. Reinke, V., Smith, H. E., Nance, J. *et al.* (2000), 'A global profile of germline gene expression in *C. elegans*', *Mol. Cell*, Vol. 6, pp. 605–616.
43. Piano, F., Schetter, A. J., Morton, D.G. *et al.* (2002), 'Gene clustering based on RNAi phenotypes of ovary-enriched genes in *C. elegans*', *Curr. Biol.*, Vol. 12, pp. 1959–1964.
44. Colaiacovo, M. P., Stanfield, G. M., Reddy, K. C. *et al.* (2002), 'A targeted RNAi screen for genes involved in chromosome morphogenesis and nuclear organization in the *Caenorhabditis elegans* germline', *Genetics*, Vol. 162, pp. 113–128.
45. Hanazawa, M., Mochii, M., Ueno, N. *et al.* (2001), 'Use of cDNA subtraction and RNA interference screens in combination reveals genes required for germ-line development in *Caenorhabditis elegans*', *Proc. Natl. Acad. Sci. USA*, Vol. 98, pp. 8686–8691.
46. Zhang, Y., Ma, C., Delohery, T. *et al.* (2002), 'Identification of genes expressed in *C. elegans* touch receptor neurons', *Nature*, Vol. 418, pp. 331–335.
47. Roy, P. J., Stuart, J. M., Lund, J. and Kim, S. K. (2002), 'Chromosomal clustering of muscle-expressed genes in *Caenorhabditis elegans*', *Nature*, Vol. 418, pp. 975–979.
48. Kim, S. K., Lund, J., Kiraly, M. *et al.* (2001), 'A gene expression map for *Caenorhabditis elegans*', *Science*, Vol. 293, pp. 2087–2092.
49. Blumenthal, T., Evans, D., Link, C. D. *et al.* (2002), 'A global analysis of *Caenorhabditis elegans* operons', *Nature,* Vol. 417, pp. 851–854.
50. GuhaThakurta, D., Palomar, L., Stormo, G.D. *et al.* (2002), 'Identification of a novel *cis*-regulatory element involved in the heat shock response in *Caenorhabditis elegans* using microarray gene expression and computational methods', *Genome Res.*, Vol. 12, pp. 701–712.
51. Gaudet, J. and Mango, S. E. (2002), 'Regulation of organogenesis by the *Caenorhabditis elegans* FoxA protein PHA-4', *Science,* Vol. 295, pp. 821–825.
52. Hill, A. A., Hunter, C. P., Tsung, B. T. *et al.* (2000), 'Genomic analysis of gene expression in *C. elegans*', *Science*, Vol. 290, pp. 809–812.
53. URL: http://www.genome.wustl.edu/projects/celegans/microarray
54. Velculescu, V. E., Zhang, L., Vogelstein, B. and Kinzler, K. W. (1995), 'Serial analysis of gene expression', *Science,* Vol. 270, pp. 484–487.
55. Pleasance, E. D., Marra, M. A. and Jones, S. J. (2003), 'Assessment of SAGE in transcript identification', *Genome Res.*, Vol. 13, pp. 1203–1215.
56. Jones, S. J., Riddle, D. L., Pouzyrev, A. T. *et al.* (2001), 'Changes in gene expression associated with developmental arrest and longevity in *Caenorhabditis elegans*', *Genome Res.*, Vol. 11, pp. 1346–1352.
57. Lynch, A. S., Briggs, D. and Hope, I. A. (1995), 'Developmental expression pattern screen for genes predicted in the *C. elegans* genome sequencing project', *Nat. Genet.*, Vol. 11, pp. 309–313.
58. McKay, S. J., Johnsen, R., Khattra, J. *et al.* (2003), 'Gene expression profiling of cells,

tissues and developmental stages of the nematode *C. elegans*', *Cold Spring Harbor Sym. Quant. Biol.*, Vol. LXVIII (in press).

59. Praitis, V., Casey, E., Collar, D. *et al.* (2001), 'Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*', *Genetics*, Vol. 157, pp. 1217–1226.
60. Fraser, A. (2003), 'Worms in L.A', *Nat. Genet.*, Vol. 35, pp. 3–5.
61. Walhout, A. J., Sordella, R., Lu, X. *et al.* (2000), 'Protein interaction mapping in *C. elegans* using proteins involved in vulval development', *Science*, Vol. 287, pp. 116–122.
62. Walhout, A. J., Temple, G. F., Brasch, M. A. *et al.* (2000), 'GATEWAY recombinational cloning: Application to the cloning of large numbers of open reading frames or ORFeomes', *Methods Enzymol.*, Vol. 328, pp. 575–592.
63. Davy, A., Bello, P., Thierry-Mieg, N. *et al.* (2001), 'A protein-protein interaction map of the *Caenorhabditis elegans* 26S proteasome', *EMBO Rep.*, Vol. 2, pp. 821–828.
64. Boulton, S. J., Gartner, A., Reboul, J. *et al.* (2002), 'Combined functional genomic maps of the *C. elegans* DNA damage response', *Science*, Vol. 295, pp. 127–131.
65. Walhout, A. J., Reboul, J., Shtanko, O. *et al.* (2002), 'Integrating interactome, phenome, and transcriptome mapping data for the *C. elegans* germline', *Curr. Biol.*, Vol. 12, pp. 1952–1958.
66. Adams, M. W. W., Dailey, H. A., DeLucas, L. J. *et al.* (2003), 'The Southeast Collaboratory for Structural Genomics: A high-throughput gene to structure factory', *Acc. Chem. Res.*, Vol. 36, pp. 191–198.
67. Mawuenyega, K. G., Kaji, H., Yamuchi, Y. *et al.* (2003), 'Large-scale identification of *Caenorhabditis elegans* proteins by multidimensional liquid chromatography-tandem mass spectrometry', *J. Proteome Res.*, Vol. 2, pp. 23–35.
68. Krijgsveld, J., Ketting, R. F., Mahmoudi, T. *et al.* (2003), 'Metabolic labeling of *C. elegans* and *D. melanogaster* for quantitative proteomics', *Nat. Biotechnol.*, Vol. 21, pp. 927–931.
69. Morino, K., Katsumi, H., Akahori, Y. *et al.* (2001), 'Antibody fusions with fluorescent proteins: a versatile reagent for profiling protein expression', *J. Immunol. Methods*, Vol. 257, pp. 175–184.
70. Stein, L. D., Bao, Z., Blasiar, D. *et al.* (2003), 'The genome sequence of *Caenorhabiditis briggsae*: A platform for comparative genomics', *PLoS Biol.*, Vol. 1, pp. 166–192.
71. Shabalina, S. A. and Kondrashov, A. S. (1999), 'Pattern of selective constraint in *C. elegans* and *C. briggsae* genomes', *Genet. Res.*, Vol. 74, pp. 23–30.
72. Kent, W. J. and Zahler, A. M. (2000), 'Conservation, regulation, synteny, and introns in a large-scale *C. briggsae–C. elegans* genomic alignment', *Genome Res.*, Vol. 10, pp. 1115–1125.
73. Coghlan, A. and Wolfe, K. H. (2002), 'Fourfold faster rate of genome rearrangement in nematodes than in *Drosophila*', *Genome Res.*, Vol. 12, pp. 857–867.
74. Lim, L. P., Lau, N. C., Weinstein, E. G. *et al.* (2003), 'The microRNAs of *Caenorhabditis elegans*', *Genes Dev.*, Vol. 17, pp. 991–1008.