Asako Sugimoto

High-throughput RNAi in Caenorhabditis elegans: genome-wide screens and functional genomics

Received November 9, 2003; accepted December 13, 2003

Abstract The phenomenon of RNA-mediated interference (RNAi) was first discovered in the nematode Caenorhabditis elegans, in which introduction of double-stranded RNA causes specific inactivation of genes with corresponding sequences. Technical advances in RNAi methodology and the availability of the complete genome sequence have enabled the high-throughput, genome-wide RNAi analysis of this organism. Several groups have used large-scale RNAi to systematically examine every C. elegans gene for knock-down phenotypes, providing basal information to be mined in more detailed studies. Now, in addition to functional genomic RNAi analyses, high-throughput RNAi is also routinely used for rapid, genome-wide screens for genes involved in specific biological processes. The integration of high-throughput RNAi experiments with other large-scale data, such as DNA microarrays and protein-protein interaction maps, enhances the speed and reliability of such screens. The accumulation of RNAi phenotype data dramatically accelerates our understanding of this organism at the genetic level.

Key words C. elegans · RNAi · functional genomics

Introduction

One of the most powerful methods in the study of gene function is the characterization of phenotypes caused by alterations in gene activity. The availability of large collections of mutations in the genome of the nematode Caenorhabditis elegans has made this organism one of the most amenable systems for such genetic analysis. Several years ago, investigators demonstrated a novel phenomenon in C. elegans in which the introduction of double-stranded RNA (dsRNA) caused the specific degradation of mRNA (Fire et al., 1998). This phenomenon, named RNA interference (RNAi), was soon recognized as an experimentally simple and technically undemanding method for gene knock-down in this organism and other species. The emergence of this new technology, which coincided with the completion of the sequencing of the C. elegans genome (The C. elegans Sequencing Consortium, 1998), has brought about a dramatic shift in the experimental strategies adopted in the study of this organism, and greatly expanded our understanding of gene function at the whole-genome level. In this review, I summarize the ways in which large-scale RNAi is being used to produce new insights in a wide variety of biological processes in C. elegans.

RNAi in C. elegans

The first report on RNA-mediated interference was published in 1995 (Guo and Kemphues, 1995), in which injection of antisense RNA corresponding to the par-1 gene caused a phenotype that mimicked that of the par-1 mutant. Surprisingly, the authors of that study found that injection of sense RNA inhibited the activity of the corresponding gene just as severely as did the antisense strand, providing an initial clue indicating that the mechanism underlying this phenomenon was something other than translational inhibition caused by annealing of the antisense nucleic acids to the mRNA. Fire et al. (1998) followed up with a systematic set of experiments and found that dsRNA (apparently produced in small amounts in the in vitro synthesis of single-stranded RNA) is a potent effector of gene interference. A series

Asako Sugimoto
Laboratory for Developmental Genomics
RIKEN Center for Developmental Biology
2-2-3, Minatojima-Minamimachi
Chuo-ku, Kobe, 650-0047, Japan
Tel: +81-78-306-3257
Fax: +81-78-306-3255
E-mail: sugimoto@cdb.riken.jp

U.S. Copyright Clearance Center Code Statement: 0301–4681/2004/7202/3–81 $15.00/0
dsRNAs are processed into "secondary siRNAs" that mRNA as templates (Sijen et al., 2001). Generated new dependent RNA polymerase (RdRP), using the target prime dsRNA synthesis through the action of RNA-genetic and biochemical analyses suggest that siRNAs the amplification of RNAi signals. Findings from systemic RNAi is defective in the C. elegans sid-1 intercellular transport of dsRNA, which is a necessary (on dsRNA length (Feinberg and Hunter, 2003). Drosophila potency of mutant, while cell-autonomous RNAi is unaffected. C. elegans membrane protein in (Winston et al., 2002). SID-1, a multispan transmembrane, has been implicated in the underlying mechanisms of which are not fully understood. In particular, genes that function in the nervous system have proven refractory to RNAi (Tavernarakis et al., 2000). In addition, there is significant inter-experimental variability in RNAi results, which may be due to subtle differences in experimental conditions, such as the developmental stage at which the worms are treated or the concentration of dsRNA (Simmer et al., 2003). Researchers should therefore interpret RNAi results carefully, taking into account the penetrance and variability of the RNAi effects.

Although this “core RNAi mechanism” appears to be conserved among diverse organisms, some aspects of RNAi observed in C. elegans are not universally found in other species. First, while long (>100 base pairs) dsRNA molecules are potent effectors of gene-specific silencing in C. elegans and other invertebrate species (Fire et al., 1998), in mammals, long dsRNA induces the activation of sequence-nonspecific dsRNA responses, such as the RNA-dependent protein kinase (PKR) pathway, that repress general translation. Because of this, siRNAs or short-hairpin RNAs (shRNAs) must be used instead of long dsRNAs to knock down genes in mammalian cells (Elbashir et al., 2001; Paddison et al., 2002).

Second, in C. elegans and some other invertebrates and plants, gene silencing is observed even in body regions remote from the site of the initial dsRNA delivery. This phenomenon is called “systemic RNAi” (Winston et al., 2002). SID-1, a multispan transmembrane protein in C. elegans, has been implicated in the intercellular transport of dsRNA, which is a necessary process in systemic RNAi (Winston et al., 2002); systemic RNAi is defective in the C. elegans sid-1 mutant, while cell-autonomous RNAi is unaffected. Furthermore, SID-1 expression sensitizes the RNAi potency of Drosophila S2 cells in a manner dependent on dsRNA length (Feinberg and Hunter, 2003). (Drosophila possesses cell-autonomous RNAi but lacks systemic RNAi, and its genome does not contain detectable sid-1 homologs.) Thus, SID-1 appears to enable the passive cellular uptake of dsRNA.

Another intriguing aspect of RNAi in C. elegans is the amplification of RNAi signals. Findings from genetic and biochemical analyses suggest that siRNAs prime dsRNA synthesis through the action of RNA-dependent RNA polymerase (RdRP), using the target mRNAs as templates (Sijen et al., 2001). Generated new dsRNAs are processed into “secondary siRNAs” that can silence the expression of the target genes. The amplification of siRNAs and their efficient intercellular transport contribute to the robustness and persistence of the RNAi effect in C. elegans.

Although RNAi in C. elegans is a powerful method for the inactivation of gene function, it does have several limitations. There are some tissue-specific and gene-specific differences in sensitivity to RNAi, the underlying mechanisms of which are not fully understood. In particular, genes that function in the nervous system have proven refractory to RNAi (Tavernarakis et al., 2000). In addition, there is significant inter-experimental variability in RNAi results, which may be due to subtle differences in experimental conditions, such as the developmental stage at which the worms are treated or the concentration of dsRNA (Simmer et al., 2003). Researchers should therefore interpret RNAi results carefully, taking into account the penetrance and variability of the RNAi effects.

Even though RNAi offers a simple and rapid method for gene inactivation, genetic mutants obtained from traditional forward genetics will remain vital for genetic analysis; a wide variety of mutant characteristics, not only loss-of-function, but also gain-of-function and temperature-sensitive mutations, will continue to provide important information regarding gene functions. In addition, the use of stable mutants is preferred to temporal RNAi knock-down for detailed genetic analyses (e.g., epistatic analysis). Given the respective strengths of the approaches, RNAi should not be viewed as a replacement for mutant-based analyses; rather, the two strategies should be regarded as complementary.

### High-throughput RNAi

Soon after the first report of the RNAi phenomenon, it became a routine method for knocking down gene function in C. elegans. It did not take long for several groups to develop methodologies to perform large-scale RNAi, thereby enabling the performance of genome-wide analyses. Two of the major factors in these scaled-up projects were the choice of dsRNA delivery method, and the sources of gene collections used in the synthesis of dsRNA.

#### Choice of dsRNA delivery methods

Four methods are available for delivering dsRNA into the C. elegans body (Fig. 1). The most potent method of delivery is microinjection, in which dsRNA is injected into the body of an adult or late larva and the RNAi phenotypes are observed in the next generation (Fig. 2A; Fire et al., 1998). In the soaking method, worms are immersed in a concentrated dsRNA solution without food for >24 hours, and recovered onto conventional
culture plates for observation of phenotypes of the soaked worms and their progeny (Tabara et al., 1998; Maeda et al., 2001). In the feeding method, dsRNA-expressing bacteria are fed to worms on agar plates (Timmons and Fire, 1998). A major advantage of the feeding method is that large numbers of worms can be treated at one time, which is of value in biochemical experiments. By allowing investigators to select the developmental stage for dsRNA delivery, both the soaking and the feeding methods can be used to conduct stage-specific RNAi experiments (Fig. 2B). In the case of the soaking method, to perform post-embryonic specific knock-downs, L1 larvae (instead of L4 or adults) are soaked in dsRNA solution (Kuroyanagi et al., 2000). While soaking (i.e., while without food), the development of L1 larvae is arrested but the RNAi response continues, making it possible to suppress gene function from the onset of post-embryonic development. In contrast, because the RNAi effect is elicited gradually during post-embryonic development even if L1 larvae are fed, the feeding method is better suited to the analysis of late larval development and of phenotypes that require persistent RNAi effect, such as analyses of behavior and longevity (Dillin et al., 2002; Lee et al., 2003; see below).

Finally, hairpin RNA that elicit RNAi response can be expressed in vivo from transgenes (Tavernarakis et al., 2000). Some neuronal genes are more effectively inactivated by transgene-based RNAi than by other delivery methods (Tavernarakis et al., 2000). The transgene-based method enables heritable, inducible, or tissue-specific RNAi by the use of specific promoters (e.g., heat-shock promoters or neuron-specific promoters).

Among these four delivery methods, the soaking and feeding methods are the most readily applicable to high-throughput analyses. Although microinjection is more labor intensive than soaking or feeding, because it causes severe gene inhibition more reliably than other methods, it is also sometimes used for large-scale analyses. However, the transgene-based method is unsuitable for high-throughput analysis, as the construction of transgenic strains requires multiple steps for each gene.

Choice of gene library
To perform genome-wide or large-scale RNAi, a gene library is required to provide the templates for dsRNA synthesis. There are two strategies for the construction of gene libraries for RNAi: use of genomic sequence information or cDNA libraries. Two groups have used genomic PCR products based on predicted gene structures to perform functional genomic RNAi analyses (see below). While T. Hyman’s group used in vitro synthesized dsRNA for injection (Gonczy et al., 2000), J. Ahringer’s group cloned the PCR products into a plasmid vector expressing dsRNA in E. coli (Fraser et al., 2000; Kamath et al., 2003). This bacterial strain library, which contains 86% of the genes predicted for C. elegans, is now available to the public.

Two other groups used cDNA collections: F. Piano’s group used an ovary-enriched cDNA library (Piano et al., 2000), whereas our group used a nonredundant cDNA set established by Y. Kohara (Maeda et al., 2001). Thus far, Kohara’s C. elegans EST project has isolated cDNAs corresponding to ~10,000 genes, representing more than half of the predicted number of genes (Reboul et al., 2001). M. Vidal’s group recently published the C. elegans “ORFeome” consisting of ~12,000 cloned ORFs (Reboul et al., 2003). This ORFeome set has not been used for genome-wide RNAi analysis, but does provide an alternative source for the in vitro synthesis of dsRNA for use for injection or soaking.

Each type of template set has its advantages and disadvantages. Although PCR product collection based on gene prediction provides a good general representation of the vast majority of genes in the genome, there are also inevitable errors and omissions. Nine percent of
known genes were not predicted by the computer analysis of the genome sequence, but were rather identified from isolated cDNAs (Reboul et al., 2001). In addition, a comparison of cDNA sequence and gene prediction using GeneFinder software revealed that more than 50% of predicted genes needed corrections to their exon-intron structures (Reboul et al., 2003). Because mature mRNAs are the target molecules in RNAi, misprediction of gene structure might reduce the efficacy of the RNAi response. cDNA sets, however, represent only actual gene expression. cDNA templates tend to be longer than PCR-based templates, which makes them more likely to produce stronger RNAi effects. However, the cDNA libraries that are currently available cover at most ~60% of the genes in the genome (Reboul et al., 2003), and do not contain genes expressed at extremely low levels. The use of these two types of gene libraries, giving consideration to their relative merits, should provide complementary coverage of the genome-wide RNAi analysis.

Large-scale RNAi analysis can be categorized into two types. The first type is the functional genomic analysis, which aims to assign in vivo gene functions to every gene by systematically recording RNAi phenotypes for each gene. The second type of large-scale RNAi involves screening for genes involved in specific developmental or biological processes. In this type of analysis, as in conventional forward genetic screens, specific assays are designed to detect abnormalities in specific phenomena. Both types of large-scale RNAi experiments are discussed below.
Functional genomics by RNAi

Several groups have performed functional genomic analyses by high-throughput RNAi (Table 1). Ahringer’s group used the feeding method to perform RNAi analysis for ~86% of predicted genes (Fraser et al., 2000; Kamath et al., 2003). Hyman’s group has been performing systematic RNAi by injection, using the PCR-amplified genomic fragments as templates, and has published the results of their analysis of chromosome III (Gonczy et al., 2000). Our group reported RNAi results for ~2,500 genes knocked down using the soaking method in which a nonredundant cDNA set was used to provide templates for dsRNA (Maeda et al., 2001). Piano and Kemphues’s group has also performed a smaller scale analysis using an ovary cDNA library (Piano et al., 2000).

In all of these studies, dissecting microscopes were used to analyze developmental abnormalities caused by RNAi for each gene. The recorded phenotypes include embryonic or larval lethality, sterility, morphological defects, locomotion defects, and retarded growth. Because neuronal cells are less sensitive to RNAi, many genes that play essential functions in neurons are likely to have been missed in these analyses. In addition, phenotypes that require specific assays (e.g., aging, stress response) and male-specific phenotypes (male-specific morphogenesis, mating behavior, etc.) were not examined, and will have to be screened separately (see below). In their secondary analyses, the groups focused on different developmental processes. Ahringer, Hyman, and Piano’s groups mainly focused on early cell divisions (Gonczy et al., 2000; Piano et al., 2000; Zipperlen et al., 2001), while our group concentrated on germline development (Maeda et al., 2001).

Recently, Plasterk’s group performed another round of genome-wide RNAi with a feeding library using an RNAi-hypersensitive mutant rrf-3 as the host strain (Simmer et al., 2003). The rrf-3 gene encodes one of the four RdRP genes, and its mutants exhibit higher RNAi sensitivity (Simmer et al., 2002). Although phenotypes for developmentally essential genes were detected at nearly equal levels in both rrf-3 and wild-type, the inhibition of genes with post-embryonic phenotypes was significantly more prevalent in rrf-3 than in the wild-type strain (Simmer et al., 2003). Another striking feature of the rrf-3 results is the high number of clones that showed slowed or arrested growth (Simmer et al., 2003). That said, some genes appear to be resistant to RNAi in rrf-3, as evidenced by the fact that some known mutant phenotypes are not reproduced (Simmer et al., 2003). In addition, it should be noted that rrf-3 itself exhibits low-penetrance embryonic lethality and sterility. This might have interfered with the detection of weak phenotypes induced by RNAi, and some of the observed phenotypes might be synthetic effects of rrf-3.

Collectively, these large-scale RNAi results have provided important insights into the genome organization of C. elegans. The percentage of genes that caused detectable RNAi phenotypes ranged between 10% and 25% in the five screens; such genes are thought to play essential roles in viability and proper development (Fraser et al., 2000; Gonczy et al., 2000; Piano et al., 2000; Maeda et al., 2001; Kamath et al., 2003). Genes that have orthologs in other organisms were found to be more likely to produce phenotypes than nonconserved genes (Fraser et al., 2000; Gonczy et al., 2000; Piano et al., 2000). In addition, genes whose cDNA has been isolated, meaning those that are relatively highly expressed during normal development, were significantly more likely to show RNAi phenotypes (Fraser et al., 2000).

Another major finding from these genome-wide RNAi analyses was that different chromosomal regions may be specialized for particular functions. Genes required for viability are significantly underrepresented on the X chromosome compared to those on the autosomes (Maeda et al., 2001; Kamath et al., 2003). In the autosomes, essential genes are enriched two-fold in the central “cluster” region of each chromosome where rates of recombination are low. Furthermore, genes that show viable post-embryonic phenotypes are present in greater frequency in the center of the X chromosome (Kamath et al., 2003). There are two possible explanations

<table>
<thead>
<tr>
<th>Paper</th>
<th>Gene set (gene number)</th>
<th>dsRNA templates</th>
<th>dsRNA delivery</th>
<th>Genes that showed detectable phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraser et al., 2000;</td>
<td>~ 86% of the predicted</td>
<td>Genomic DNA (Feeding bacteria library)</td>
<td>Feeding</td>
<td>10.3%</td>
</tr>
<tr>
<td>Kamath et al., 2003</td>
<td>(16,401)</td>
<td>Genomic DNA (Feeding bacteria library)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simmer et al., 2003</td>
<td></td>
<td>Genomic DNA (PCR fragments)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gonczy et al., 2000</td>
<td>Chromosome III (2232)</td>
<td>cDNA</td>
<td>Soaking</td>
<td>27.2%</td>
</tr>
<tr>
<td>Maeda et al., 2001</td>
<td>Mixed stage nonredundant cDNA library (2479)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Piano et al., 2000</td>
<td>Ovary-derived cDNA library (350)</td>
<td></td>
<td>Injection</td>
<td>&gt; 23%</td>
</tr>
</tbody>
</table>
for this underrepresentation of essential genes on the X chromosome. It may be that mutations in the essential genes on the X chromosome directly lead to lethality or sterility in XO males, with the result that essential genes tend to be removed from the X chromosome. It has also been reported that the X chromosome is transcriptionally silenced in the germline during mitosis and early meiosis (Kelly et al., 2002), which would also contribute to a reduced incidence of genes required in germline development and early embryogenesis on this chromosome.

For the most part, the results of large-scale RNAi conducted by the various groups are consistent with each other, but there has also been some lack of agreement. This is partly because each of the laboratories used a different RNAi protocol and different definitions of phenotypes. In addition, inter-experimental comparison of the large-scale feeding RNAi data set within the same laboratory following the same experimental procedures showed 10%–30% variability in RNAi results, mainly due to the high frequency of false negatives in the RNAi screens (Simmer et al., 2003). These variable results with respect to identical genes may occur as a result of slight differences in the timing (in terms of developmental stage) of the exposure to dsRNA, changes in temperature during the experiment, or concentration of the dsRNA used. It is therefore important to take into account the variability in RNAi effect that inevitably occurs when interpreting high-throughput RNAi data, and multiple rounds of large-scale analyses will be helpful in validating the function of individual genes. (A detailed comparison of the large-scale data set has been reported elsewhere; Piano and Gunsalus, 2002.) Large-scale RNAi data sets from each group are accessible via their original databases or via WormBase (Stein et al., 2001), the integrated database of the C. elegans research community.

**RNAi-based gene screenings**

The functional genomic analyses described above focused mainly on general developmental processes. This focus meant that phenotypes involving behavior, aging, or other subtle manifestations were not detected in these screens. Many researchers now routinely perform large-scale RNAi screens to identify the genes involved in specific processes. One major advantage of the RNAi-based screen over the conventional forward genetic screen is its speed. By RNAi, the connection between gene identity and phenotype can be made immediately, whereas it may take from several months to years to clone the gene responsible for a mutant phenotype. In addition, while RNAi screens can target every gene systematically, the hit rate of chemical mutagens is uneven; large genes are statistically more prone to mutation than small genes, making mutations in small genes more difficult to isolate. Examples of successful RNAi-based gene screens using diverse strategies are given in Table 2.

Screens for genes involved in specific biological processes

The feeding library used in the functional genomic analysis has also been used in various screens for genes involved in specific processes (Dillin et al., 2002; Ashrafi et al., 2003; Lee et al., 2003; Pothof et al., 2003; Vastenhouw et al., 2003). The feeding method is suitable to assay phenotypes that appear in adulthood, as RNAi response can be elicited by the continuous feeding of dsRNA-expressing bacteria to the worms (Timmons and Fire, 1998; Kamath et al., 2001). Two groups took advantage of this method to screen for genes involved in longevity, and both found that genes important for mitochondrial function affect C. elegans lifespan (Dillin et al., 2002; Lee et al., 2003).

The development of a sensitive assay strategy is essential to achieving an effective screen. To identify genes necessary for fat storage, Ashrafi et al. (2003) used the vital dye Nile Red to visualize fat storage droplets in living worms. In a genome-wide screen using the feeding library, they identified 305 genes whose inactivation results in reduced fat storage, and 112 genes whose inactivation causes increased fat storage (Ashrafi et al., 2003). Pothof et al. (2003) used a sophisticated assay to detect increased DNA instability in which frameshifts and small insertions or deletions in somatic cells were detected by LacZ reporter expression. They performed this assay in animals treated by feeding RNAi and identified 61 genes whose knock-down increased genomic instability (Pothof et al., 2003).

Dudley et al. (2002) reported an intriguing screen in which RNAi was used to screen for genes required for RNAi, an approach the authors named “RNAi-to-RNAi.” In a screen aimed to identify genes with essential roles in embryogenesis, they serendipitously identified a dsRNA that acts as a potent suppressor of RNAi. When this dsRNA, which corresponds to the gfl-1 (human GAS41-like) gene, is co-injected with dsRNAs that normally induce lethality, only viable progeny were obtained (Dudley et al., 2002). Using this RNAi-to-RNAi assay against several other candidate genes, they also found that zfp-1, mes-3, mes-4, and mes-6 are necessary for RNAi (Dudley et al., 2002). As this screen was relatively small scale, we can expect that a further large-scale screen is likely to lead to the isolation of even more genes involved in the RNAi mechanism.

Another gene screen targeting transposon silencing also detected a gene involved in the RNAi mechanism (Vastenhouw et al., 2003). The C. elegans genome contains ~30 copies of the Tc1 transposon, which do
<table>
<thead>
<tr>
<th>Paper</th>
<th>Screen for genes involved in</th>
<th>Assay</th>
<th>Gene pool selection</th>
<th>RNAi method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Genome-wide screen by specific assays</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dillin et al., 2002; Lee et al., 2003</td>
<td>Longevity</td>
<td>Increased lifespan</td>
<td>(Chromosome I feeding library)</td>
<td>Feeding</td>
</tr>
<tr>
<td>Ashrafi et al., 2003</td>
<td>Fat regulation</td>
<td>Visualization of fat storage by Nile Red</td>
<td>(Chromosome I and II feeding library)</td>
<td>Feeding</td>
</tr>
<tr>
<td>Pothof et al., 2003</td>
<td>Genome stability</td>
<td>Reporter assay that detects frameshifts and small insertions/deletions</td>
<td>(Genome-wide feeding library)</td>
<td>Feeding</td>
</tr>
<tr>
<td>Dudley et al., 2002</td>
<td>RNAi machinery</td>
<td>RNAi resistance</td>
<td>(PCR-amplified genomic DNA of predicted genes)</td>
<td>Injection of pools of dsRNA</td>
</tr>
<tr>
<td>Vastenhouw et al., 2003</td>
<td>Transposon silencing</td>
<td>Increased frequency of reversion of a transposon-induced mutant phenotype</td>
<td>(Genome-wide feeding library)</td>
<td>Feeding</td>
</tr>
<tr>
<td><strong>Focus on specific gene families</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parrish and Xue, 2003</td>
<td>Apoptotic DNA degradation</td>
<td>TUNEL assay</td>
<td>Protein families (DNase, RNase, cyclophilins, and topoisomerases)</td>
<td>Feeding</td>
</tr>
<tr>
<td>Keating et al., 2003</td>
<td>Receptor for neurotransmitters or neuropeptides</td>
<td>Quantitative behavioral assay</td>
<td>Protein family (G protein-coupled receptors)</td>
<td>Feeding (wild-type and <em>rrf-3</em>)</td>
</tr>
<tr>
<td><strong>Combination with transcriptome data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hanazawa et al., 2001</td>
<td>Germine development</td>
<td>Nomarski microscopy and DAPI staining</td>
<td>cDNA subtraction and differential hybridization</td>
<td>Injection</td>
</tr>
<tr>
<td>Piano et al., 2002</td>
<td>Early embryogenesis</td>
<td>Time-lapse analysis</td>
<td>Microarray (ovary-enriched genes)</td>
<td>Injection</td>
</tr>
<tr>
<td>Colaiácovo et al., 2002</td>
<td>Chromosome morphogenesis</td>
<td>Cytological analysis</td>
<td>Microarray (germline-enriched)</td>
<td>Injection</td>
</tr>
<tr>
<td>Murphy et al., 2003</td>
<td>Longevity (Downstream of DAF-16)</td>
<td>Lifespan</td>
<td>Microarray (genes whose expression is affected by DAF-16 pathway)</td>
<td>Feeding</td>
</tr>
<tr>
<td><strong>Combination with interactome data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boulton et al., 2002</td>
<td>DNA damage response</td>
<td>γ-irradiation response</td>
<td>Protein-protein interaction mapping (yeast 2-hybrid screen)</td>
<td>Feeding</td>
</tr>
<tr>
<td><strong>Use of functional genomic RNAi data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zipperlen et al., 2001</td>
<td>Early embryogenesis</td>
<td>Video-microscopy</td>
<td>Functional genomic RNAi (embryonic lethal genes)</td>
<td>Feeding and injection</td>
</tr>
</tbody>
</table>
not jump in the germline because the transposition is “silenced.” It has been known that this transposon silencing mechanism shares components with both the RNAi machinery and transgene-induced co-suppression (Ketting et al., 1999; Tabara et al., 1999). Vastenhouw et al. (2003) screened for genes involved in transposon silencing by feeding RNAi. The screen revealed increased transposition activity by the phenotypic reversion of a transposon-induced unc-22 mutant allele, and resulted in the isolation of 27 genes (Vastenhouw et al., 2003), including mut-16, a mutator gene that had previously been isolated in mutants showing defects in transposon silencing, RNAi, and transgene co-suppression (Vastenhouw et al., 2003). Three other genes, including ppw-2, a member of the argonaute family, were also shown to be involved in both transposon silencing and co-suppression (Vastenhouw et al., 2003). Although this screen did isolate a number of interesting new genes, the fact that some known mutator genes were missed suggests that the gene list has yet to be completed.

Specific gene families

In addition to large-scale studies in which RNAi is performed on every gene in the genome, it is also possible to select subsets of genes for RNAi analysis. The studies described below targeted particular gene families and assessed their in vivo function by specific assays.

Parrish and Xue (2003) selected a group of gene families (DNase, RNase, cyclophilins, and topoisomerases) as candidates in a study designed to identify genes involved in apoptotic DNA degradation. To detect abnormal apoptotic DNA degradation, they used a TUNEL assay to label the ends of intermediate chromosome fragments during apoptosis. In the nuc-1 or cps-6 mutants, both of which are known to mediate apoptotic DNA degradation, TUNEL-positive nuclei have been shown to accumulate. They performed RNAi by injection on the 77 candidate nuclease genes, and identified seven nuclease genes whose loss of function increased the number of TUNEL-positive nuclei, as in the nuc-1 and cps-6 mutants (Parrish and Xue, 2003). Further epistatic analysis utilizing RNAi indicated that there are at least two independent pathways for apoptotic DNA degradation (Parrish and Xue, 2003).

Keating et al. (2003) focused on evolutionarily conserved rhodopsin-like G protein-coupled receptors (GPCRs), and assessed their in vivo functions in a feeding RNAi study using both wild-type and RNAi-hypersensitive rrf-3 mutant strains. Working from the hypothesis that GPCRs bind either small-molecule neurotransmitters or neuropeptides and are thus involved in neural circuitry, they assayed for defects in locomotion and the rate of egg laying. Among the 60 genes tested, RNAi of seven genes resulted in uncoordinated locomotion, and six in altered egg-laying rates (Keating et al., 2003). As neuronal tissues are less susceptible to RNAi, it is conceivable that some GPCR genes might have been missed in this screen. Despite the expectation that the detection rate would increase when the rrf-3 mutant strain was used, the authors did not identify any genes in addition to those found in wild-type experiments (Keating et al., 2003). Because their assay was not designed to detect other types of neuronal defects (such as chemosensation, olfaction, or learning), it is highly plausible that some GPCR genes without RNAi phenotypes in this report nonetheless have some function in neural circuits.

Combination with transcriptome data

Use of large-scale transcriptional profile data can also be used to pre-screen genes to be tested in an RNAi screen.

Hanazawa et al. (2001) used cDNA subtraction and differential hybridization to isolate genes involved in germline development. They selected 168 genes whose mRNAs are enriched in the germline, and found that RNAi of 15 of these genes caused sterility due to a variety of defects in germline development (Hanazawa et al., 2001). The fraction of genes causing germ-line-specific defects in their screen was 9%, compared to ~1% in the genome-wide screen, indicating the enrichment procedure’s efficacy.

More recently, DNA microarrays have provided genome-wide expression profile data that can be used to narrow down the list of genes to be tested by RNAi. To identify genes involved in early embryogenesis, Piano et al. used DNA microarray data from Reinke et al. to identify germline-enriched genes (Piano et al., 2002; Reinke et al., 2000). With the expectation that mRNAs stored in oocytes would be enriched in genes involved in early cell divisions, Piano et al. (2002) selected 766 genes highly expressed in ovary, and characterized their RNAi phenotypes by time-lapse recording. As expected, the RNAi of as many as 43% of these genes caused defects in egg production or embryogenesis (Piano et al., 2002), in contrast to less than 10% in the genome-wide screens. Piano et al. (2002) also developed a systematic way to describe early embryonic phenotypes using 47 discrete phenotypic characters amenable to computer-assisted analyses. Using these digitized “phenotypic signatures,” they were able to cluster genes by phenotypic similarity. This introduction of this type of structured format for describing phenotypes will make the burgeoning RNAi data set easier to share and use.

The same RNAi data were used by Reinke et al. (2000) to identify genes involved in chromosome morphogenesis and nuclear organization during meiotic prophase. Colaiácovo et al. (2002) selected 192 germ-
line-enriched genes whose expression profiles were similar to those of previously identified meiosis genes. They performed detailed cytological analysis of the RNAi-induced phenotypes of these genes, and detected strong germline phenotypes for 27% of the genes tested (Colaiacovo et al., 2002). From this analysis, they identified genes involved in meiotic events, germline proliferation, and chromosome organization/segregation.

Murphy et al. (2003) used microarray analysis to identify genes downstream of DAF-16, a FOXO family transcription factor that influences the rate of aging in response to insulin-like growth factor signaling. Based on the microarray data, they selected genes either positively or negatively regulated by the DAF-16 pathway, and evaluated the effects of the selected 58 genes on lifespan using the feeding method of RNAi (Murphy et al., 2003). They found that RNAi of stress-response genes (e.g., catalase genes, a glutathione-S-transferase gene, small heat-shock protein genes) and of genes involved in antimicrobial response led to reductions in lifespan. They also found genes whose RNAi causes extended lifespan. These genes have relatively small individual effects on lifespan, meaning it would have been difficult to identify any of them in a conventional forward genetic screen. This study demonstrated how the selection of candidate genes by microarray and targeted RNAi analysis can serve as a powerful approach to the characterization of complex biological processes.

Combination with interactome data

Boulton et al. (2002) reported a study designed to identify genes involved in DNA damage response (DDR), in which a protein-protein interaction data set (interactome) obtained by high-throughput yeast two-hybrid assay was combined with RNAi analysis. Seventy-five C. elegans genes were selected based on sequence similarities to known DDR genes in other organisms, and proteome-wide two-hybrid screens were performed using the 67 DDR candidate genes as baits (eight genes were not amenable to two-hybrid screens due to their self-reactivity). Forty-five of the genes screened yielded a total of 165 interactors, of which 125 were novel proteins. To validate the involvement of these DDR orthologs and their potential interactors in DDR processes in vivo, the authors performed RNAi by feeding for each gene and examined the DDR-related phenotypes in response to γ-irradiation. From these analyses, they identified 23 genes (12 DDR orthologs and 11 novel genes) that showed detectable DDR phenotypes (Boulton et al., 2002). This study demonstrates that the combination of functional genomic approaches can facilitate the identification of novel components of an evolutionarily conserved pathway.

Use of functional genomic RNAi data

The accumulated functional genomic RNAi data are the useful source for the pre-selection of genes for further in-depth analyses. Zipperlen et al. (2001) used their chromosome I RNAi data as the first screen to select genes required for embryonic development, and selected 147 genes that showed embryonic lethal phenotypes at a high penetrance. The first three cell divisions and the terminal phenotypes of embryos RNAi treated for these genes were recorded using four-dimensional time-lapse video recording. The first screen was performed using only the feeding method, but for the video recording they used injection for RNAi delivery as well, as this method tends to elicit stronger loss-of-function phenotypes. Among the genes examined, over half showed defects in early cell divisions.

Because the data from large-scale functional genomic RNAi analyses are available and searchable from WormBase and other internet-based resources, researchers can now perform the first round of gene screening for particular RNAi phenotypes “in silico,” before beginning actual work at the bench.

Conclusions

Large-scale RNAi analyses have dramatically increased the amount of functional information for each gene of the C. elegans genome. In conjunction with other genome-wide data, such as expression profiles and protein-protein interaction maps, the RNAi phenotype data set promises to provide a rich source of information for use in the deciphering of C. elegans biology. In order to make it possible to integrate large-scale RNAi phenotype data with other genome-wide data sets, each RNAi phenotype needs to be described in a manner that can be easily converted into a format suitable for computer-assisted analysis. As Piano et al. (2002) demonstrated in their study of early embryogenesis, the development of a clearly defined vocabulary for the systematic description of phenotypes is required. Our group now aims to establish a methodology for the profiling of embryonic and post-embryonic phenotypes.

Considering that over 40% of C. elegans genes have sequence homologies with genes in other organisms (The C. elegans Sequencing Consortium, 1998), these functional genomic resources are useful not only to the C. elegans research community but to those studying other organisms as well. With the recent dramatic improvements in RNAi technology, it seems inevitable that large-scale RNAi will be applied to other model organisms in the very near future.

Acknowledgments We thank Douglas Sipp for comments on the manuscript; members of the Sugimoto lab for helpful discussions; and Naoko Sakamoto for the figures.
References


