

C. elegans

Methods and Applications

Edited by

Kevin Strange

Comparative Genomics in *C. elegans*, *C. briggsae*, and Other *Caenorhabditis* Species

Avril Coghlan, Jason E. Stajich, and Todd W. Harris

Summary

The genome of the nematode *Caenorhabditis elegans* was the first animal genome sequenced. Subsequent sequencing of the *Caenorhabditis briggsae* genome enabled a comparison of the genomes of two nematode species. In this chapter, we describe the methods that we used to compare the *C. elegans* genome to that of *C. briggsae*. We discuss how these methods could be developed to compare the *C. elegans* and *C. briggsae* genomes to those of *Caenorhabditis remanei*, *C. n. sp.* represented by strains PB2801 and CB5161, among others (**1**), and *Caenorhabditis japonica*, which are currently being sequenced.

Key Words: Nematode genomes; *Caenorhabditis* genomes; comparative genomics; genome evolution; nematodes.

1. Introduction

The *Caenorhabditis elegans* and *Caenorhabditis briggsae* genomes were the first pair of genomes from the same animal genus to be sequenced (**1a,2**). Comparison of their gene content and chromosomal structure has provided insight into the evolution of nematodes. In the last 2 yr, methods have been developed that exploit comparisons between multiple species to predict genes and regulatory elements, and to study genome evolution. We discuss how these methods could be used to compare the *C. elegans* and *C. briggsae* genomes to those of *Caenorhabditis remanei*, *C. sp.* PB2801, and *Caenorhabditis japonica*, which are currently being sequenced (**Fig. 1**; **ref. 3**).

2. Predicting Genes and Comparing Gene Structure in Related Genomes

2.1. Predicting Genes in Related Genomes

There were 19,735 protein-coding genes annotated in the *C. elegans* genome in a recent release of WormBase (WS140 [**4**]). These annotations have been

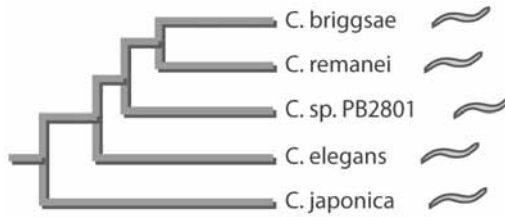


Fig. 1. The phylogenetic relationship of *Caenorhabditis elegans* and *Caenorhabditis briggsae* to *Caenorhabditis remanei*, *C. sp. PB2801*, and *Caenorhabditis japonica*, whose genomes are being sequenced. This figure is based on the phylogeny found by Kiontke et al. (19).

extensively tested and improved using computer programs and experiments. In fact, more than 17,000 *C. elegans* genes have been partially or fully confirmed by mRNA and other experiments (5). The *C. briggsae* genome is similar in size to that of *C. elegans*, approx 100 Mb, and has a similar number of genes, approx 20,000 genes (2). It will be interesting to see whether *C. remanei*, *C. sp. PB2801*, and *C. japonica* also have a similar number of genes and genome size.

The most common programs used to predict protein-coding genes are those that predict genes based on sequence alone, known as “*de novo* gene predictors.” Two *de novo* gene predictors that have been tuned for nematode genomes are Genefinder (P. Green, unpublished) and Fgenesh (6), which were used to predict genes in the *C. briggsae* genome (2). Both Genefinder and Fgenesh are relatively accurate. Genefinder predicts 48% of known *C. elegans* genes and 81% of known exons exactly right, whereas Fgenesh predicts 51% of known *C. elegans* genes and 88% of known exons correctly (7,8).

An advance in accuracy has been enabled by the development of gene predictors that use information from the alignment of the target genome to a second genome (9). These exploit the fact that protein-coding exons are better conserved over evolutionary time than nonfunctional intergenic or intronic regions. One program that uses this approach is TWINSKAN (10), which was used to predict genes in the *C. briggsae* genome using alignments to *C. elegans* (2). TWINSKAN is more accurate than Fgenesh or Genefinder. Using alignments to the *C. briggsae* genome, TWINSKAN predicts 60–63% of known *C. elegans* genes correctly, and 86–90% of known exons (8). This is partly because of the use of cross-species alignments, but is also because of more accurate prediction of intron length and of introns with noncanonical splice sites (8). The optimal evolutionary divergence for TWINSKAN is approx 90–95%, which is less than the divergence between *C. elegans* and *C. briggsae* (8). Thus, TWINSKAN will probably produce more accurate predictions for less divergent pairs of nematodes,

such as *C. briggsae* and *C. remanei* (Fig. 1). Gene predictors that use alignments between more than two species have begun to appear. Multispecies alignments are used by the program EXONIPHY to predict exons (11), and by Evogene to predict genes (12). These programs will be useful for producing accurate gene predictions for *C. remanei*, *C. sp.* PB2801, and *C. japonica*, as well as for improving *C. elegans* and *C. briggsae* annotation.

Wei et al. (8) recently demonstrated the potential of interspecies comparisons for improving *C. elegans* annotation. They used TWINSKAN to predict 265 open reading frames in the *C. elegans* genome that did not overlap existing gene predictions. They successfully cloned and sequenced 146 of these open reading frames, thereby finding evidence for 146 previously unrecognized genes. Interspecies comparisons can also identify unrecognized exons that are part of known genes. For example, by comparing the *C. elegans* epidermal growth gene *lin-3* to its *C. briggsae* ortholog, Liu et al. (13) predicted an alternatively spliced 5' exon, which they confirmed experimentally. By aligning *C. elegans* genes to the *C. briggsae* genome sequence, Stein et al. (2) suggested that there may be more than 1200 unrecognized exons in known *C. elegans* genes. Multispecies comparisons will help to distinguish how many of these putative exons are real. If a putative exon seems to be conserved in *C. remanei*, *C. sp.* PB2801, and *C. japonica*, as well as in *C. briggsae* and *C. elegans*, it is likely to be real. One caveat to remember is that some gene structures have probably changed since the species diverged. For example, changes in splice sites have led to different alternative splicing of the CPEB genes *fog-1* and *cpb-2* in the five *Caenorhabditis* species (14).

A major challenge in gene prediction is annotating species-specific genes, because we cannot use alignments between species to predict them, or to support predictions. Stein et al. (2) predicted approx 1000 *C. elegans*-specific genes and approx 110 *C. elegans*-specific gene families. Some of these may simply be genes that have evolved very rapidly. If this is true, it may be possible to identify their orthologs in other species by using synteny information, or by using sensitive homology search methods, such as profile-hidden Markov models (15). However, some of the putative species-specific genes are probably novel genes that have been generated since the species diverged (16). Even though such genes are difficult to predict, they are of great interest because they may be involved in species-specific adaptations (17).

2.2. Comparing Gene Structure in Related Genomes

The five *Caenorhabditis* genomes will be a treasure trove for studying the evolution of gene structure. There are many unsolved questions about the evolution of intron–exon structure and splicing patterns. For example, it is not known how introns are gained or lost (14,18). The *Caenorhabditis* genomes

will provide fertile ground for studying this question, because approx 9% of *C. elegans* and *C. briggsae* introns are species-specific (2). To distinguish whether the *C. elegans*-specific introns were gained by *C. elegans* or lost from *C. briggsae*, an outgroup species is needed. The *C. japonica* genome will provide such an outgroup (Fig. 1). For example, Kiontke et al. (19) cloned the gene for the largest subunit of RNA polymerase II from 10 *Caenorhabditis* species. By comparing their intron–exon structures, they estimated that there have been up to 19 intron losses and up to 15 gains in this gene since the species diverged. Similarly, Cho et al. (14) cloned five genes from six *Caenorhabditis* species, and also found loss to be more common than gain. It will be exciting to extend these studies to all 20,000 genes in the *Caenorhabditis* genomes. In particular, examination of introns that have been lost or gained very recently—those that are specific to just one species or strain—may help uncover the mechanisms of intron loss and gain. It may also reveal whether losses and gains affect gene expression or function, and could even be adaptive (20).

3. Studying Orthologs, Paralogs, and Gene Families

3.1. Identifying Orthologs and Paralogs Between Genomes

Orthologs are genes in different species that evolved from a common ancestral gene by speciation. In contrast, paralogs are genes that originated by duplication within a genome. Stein et al. (2) identified 11,255 *C. briggsae*–*C. elegans* one-to-one orthologs, by identifying *C. briggsae*–*C. elegans* gene pairs that were each other's top BLASTP match (21). To avoid contaminating the ortholog set with paralogs, ortholog pairs also had to have BLASTP *E*-values more than 10^5 lower (more significant) than the next-best match. This approach can miss one-to-one orthologs that have evolved rapidly or that belong to gene families, but ensures a high-confidence set of ortholog pairs. Stein et al. (2) identified an additional 900 ortholog pairs using conserved gene order. The final set of 12,155 one-to-one orthologs included approx 60–65% of the *C. elegans* and *C. briggsae* gene sets. The remaining 35–40% of the genes in each species are species-specific genes, or have multiple orthologs in the other species.

Additional methods to identify orthologs include InParanoid (22), which identifies one-to-one, one-to-many, and many-to-many orthologs between two genomes based on BLASTP similarity scores, providing a confidence measure for each ortholog assignment. InParanoid identifies 12,858 *C. briggsae*–*C. elegans* ortholog groups (see <http://inparanoid.cgb.ki.se/>). On the other hand, OrthoMCL (23) identifies one-to-one, one-to-many, and many-to-many orthologs between multiple genomes, by using the Markov Cluster Algorithm (MCL) (24), which groups related sequences into clusters. OrthoMCL distinguishes recent paralogs from orthologs by identifying sequences that have closer BLASTP matches within a species than between species.

The most accurate method of identifying one-to-one, one-to-many, and many-to-many orthologs between multiple genomes is to build phylogenetic trees (25). By building a phylogenetic tree of the *sra* chemosensory receptor family, Stein et al. (2) showed that assigning one-to-one orthologs using mutual best BLASTP matches leads to false-positive and false-negative assignments approx 15% of the time. However, Stein et al. could not use phylogenetic trees to identify all *C. briggsae*–*C. elegans* orthologs, because they lacked sequences from an outgroup to root the trees. The *C. japonica* genome will provide an outgroup (Fig. 1), and so will allow *C. briggsae*–*C. elegans* orthologs to be identified accurately.

3.2. Identifying Gene Families

Gene duplication provides a substrate for new gene functions (26), so that lineage-specific duplications often trigger lineage-specific adaptations. In *C. elegans* several gene families have undergone lineage-specific expansions, including subfamilies of the seven-transmembrane (7TM) proteins, some of which function as chemoreceptors (27). These 7TM gene duplications occurred in the *C. elegans* genome after divergence from *C. briggsae*, in two specific chemoreceptor families (2,28).

The TRIBE-MCL program can be used to identify gene families (29). It builds a graph of all proteins based on sequence similarity, where each protein is a node in the graph, and the distance between nodes is a function of the similarity between proteins. Most pairs of nodes are not connected by edges, so the graph is sparse. The TRIBE-MCL program finds clusters of similar nodes based on the internode distances, using the MCL (24). MCL is an algorithm that applies a series of expansions and contractions to the graph in order to identify robust clusters. MCL differs from simpler clustering methods, such as single-linkage clustering, which builds a cluster of proteins by finding all interconnected nodes. A limitation of single-linkage clustering is its tendency to build clusters that are too heterogeneous, because a single protein domain shared among otherwise unrelated proteins serves as a connector between all the nodes.

To identify gene families that are significantly larger in one species than another, TRIBE-MCL can be used to build gene families using the entire protein sets from the two species. A χ^2 test can then be used to identify families that have a significantly greater copy number in one species than in the other. Applying this approach to multiple nematode genomes will allow us to identify families that have contracted or expanded within each lineage. For example, in the genomes of parasitic nematodes such as *Brugia malayi* (30) one might expect gene families necessary for a free-living lifestyle to have contracted and families related to evasion of the host immune system to have expanded.

Comparison of the *C. elegans* and *C. briggsae* genomes revealed that several families have significantly expanded within each lineage (2). Table 1 lists the 26

Table 1
The 26 Largest Gene Families in *Caenorhabditis elegans* and *Caenorhabditis briggsae*^a

Cluster	<i>C. briggsae</i>	<i>C. elegans</i>	Result of χ^2 test	Cluster function
0	299	275	Not significant	Protein kinase
1	268	284	Not significant	Transcription factor, zinc finger
2	224	266	Not significant	7TM subfamily 1
3	163	312	Significant	7TM subfamily 2
4	230	184	Not significant	EGF-like
5	222	19	Significant	Unknown function
6	122	116	Not significant	BTB/POZ
7	97	141	Significant	Serpentine receptor (some are pseudogenes)
8	14	211	Significant	F-box/FTH
9	88	137	Significant	Lectin
10	93	111	Not significant	Unknown: DUF216
11	102	96	Not significant	Cuticle collagen
12	80	117	Not significant	Tyrosine specific protein phosphatase
13	103	94	Not significant	Neurotransmitter-gated ion-channel ligand binding domain
14	97	89	Not significant	Myosin
15	94	92	Not significant	G protein β WD-40 repeat
16	91	90	Not significant	UDP-glucuronosyl/UDP-glucosyl transferase
17	87	94	Not significant	Protein kinase
18	96	80	Not significant	Transcription factor, zinc finger
19	88	79	Not significant	RNA binding protein
20	73	82	Not significant	Cytochrome P450
21	78	70	Not significant	Cuticle collagen
22	72	70	Not significant	Ras GTPase
23	62	80	Not significant	Lectin
24	74	67	Not significant	Rhodopsin-like GPCR superfamily (serpentine receptor)
25	10	130	Significant	F-box

^aFound using TRIBE-MCL (29). For each family, a χ^2 test was used to test whether the size of the family is significantly different in the two species.

7TM, seven transmembrane; EGF, epidermal growth factor; BTP, bric-a-brac, tramtrack, and Broad-Complex proteins; POZ, Pox virus and Zinc finger; FTH, FOG-2 homology domain; WD, G- β repeat domain (Trp-Asp dipeptide); GPCR, G protein-coupled receptor.

largest families in each species. One of the most striking examples of *C. elegans*-specific expansion is a subfamily of 7TM chemoreceptors. A detailed analysis of this family showed that the expansion was caused by a series of local tandem duplications, presumably through unequal crossing over (28). Many of the other apparent gene family expansions were actually owing to duplications of pseudogenes, rather than to duplication of functional genes. For example, cluster 7 in Table 1 initially appeared to be because of an expansion of serpentine and rhodopsin-like G protein-coupled receptor genes, but many of these genes have since been identified as pseudogenes. Identifying significantly expanded families in the *C. elegans* and *C. briggsae* lineages is a very active area of research. It is not known how important these expansions are to the current lifestyle of each species. For example, it will be important to determine whether recently acquired chemoreceptors provide additional sensing for *C. elegans*, perhaps by binding additional ligands or by using combinations of receptors.

To more accurately identify families that have expanded recently, sophisticated models of gene family evolution are now being developed, incorporating the time since speciation and the rate of gene duplication. The best approach for identifying lineage-specific expansions is to use phylogenetic trees, including family members from more than two species (Hahn, M. W., De Bie, T., Stajich, J. E., Nguyen, C., and Cristianini, N., manuscript in prep.).

4. Computational Prediction of Regulatory Elements

The discovery of sequence elements that control the temporal and spatial expression of genes is a difficult but exciting area of comparative genomics. The small size of intergenic regions in *C. elegans* and *C. briggsae*, which are generally less than 1.5 kb, helps in this search. Furthermore, *C. elegans* and *C. briggsae* seem to be at an ideal evolutionary distance for identifying functionally important noncoding sequences. Webb et al. (31) compared 142 orthologous intergenic regions in *C. elegans* and *C. briggsae*, and found a mosaic pattern with regions of high similarity scattered among longer nonalignable regions. They suggested that the regions of high similarity might contain regulatory sequences.

A standard approach for finding candidate regulatory sequences is to align 0.5–2.0 kb upstream of orthologous *C. elegans* and *C. briggsae* genes. An important step is to repeat-mask the sequences first—to reduce the amount of spurious alignments. The upstream sequences can then be aligned using a local alignment algorithm, such as BLASTZ (32), or a global alignment algorithm, such as GLASS (33), or both. The alignment is scanned for conserved noncoding sequences (CNSs): long stretches of high-sequence identity, for example, of more than 70% identity over more than 50 bp. This approach has been successfully used to find regulatory sequences that control *C. elegans* pharyngeal development (Fig. 2; ref. 34) and vulval expression (35).

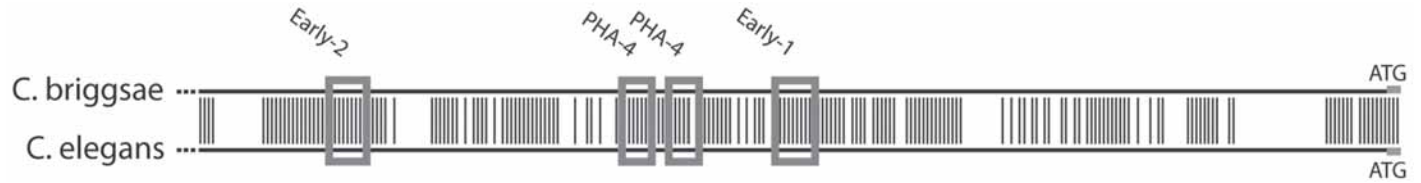


Fig. 2. A conserved noncoding sequence (CNS) found by Gaudet et al. (34) upstream of the *Caenorhabditis elegans* pharyngeal gene *K07C11.4* and its *Caenorhabditis briggsae* ortholog. Gaudet et al. identified four conserved transcription factor binding sites in this CNS (shown as gray boxes).

Identifying transcription factor binding sites within CNSs is difficult, because they can be as short as 6 bp. A common approach is to search for matches to known transcription factor binding sites in the TRANSFAC database (36). Alternatively, if several *C. elegans* genes are probably regulated by the same transcription factor, one can search for short motifs that are overrepresented in CNSs in the upstream regions of the *C. elegans* genes and their *C. briggsae* orthologs (34,35,37). This approach has the advantage that it can identify novel binding sites that are absent from TRANSFAC. For example, Gaudet et al. (34) discovered novel transcription factor binding sites in *C. elegans* genes expressed in the pharynx, by using the program ImProbizer (38) to find motifs that are unexpectedly frequent in their upstream regions (Fig. 2). Many other programs exist for identifying potential transcription factor binding sites in CNSs. Tompa et al. (39) recently assessed the accuracy of 13 motif-discovery programs in identifying known transcription factor binding sites. One of the most successful programs was Weeder (40), which counts how often all possible motifs up to a certain length occur in the input sequences, and identifies overrepresented motifs that are well conserved.

GuhaThakurta et al. (37) tested the hypothesis that most *C. elegans* regulatory sequences can be found by searching within CNSs in orthologous *C. briggsae*–*C. elegans* upstream regions. They aligned 2 kb upstream of 33 *C. elegans* muscle-expressed genes to the upstream regions of their *C. briggsae* orthologs, using both local and global alignment algorithms. They then scanned the alignments to find CNSs with more than 65% identity over more than 50 bp. For some *C. elegans* genes the CNSs contained 55% of predicted muscle transcription factor binding sites, but for other genes they contained just 6% of predicted sites. GuhaThakurta et al. suggested that errors in the alignments used to identify CNSs might explain why CNSs do not contain all *C. elegans* transcription factor binding sites. Another reason is that the number of binding sites, their positions and sequences probably differ between *C. elegans* and *C. briggsae*. For example, the *lin-48* gene has a conserved expression pattern in *C. elegans* and *C. briggsae*, but the *C. elegans* upstream region contains two binding sites for the transcription factor EGL-38, whereas the *C. briggsae* promoter contains just one site (41). Furthermore, the *C. elegans* and *C. briggsae* EGL-38 binding sites have diverged in sequence, so that the *C. briggsae* site cannot compensate for mutations in the *C. elegans* sites.

Comparison of the upstream regions of the *C. elegans* and *C. briggsae* *lin-3* genes to those of their *C. remanei* and *C. sp. PB2801* orthologs suggests that multi-species comparisons will help to distinguish functionally important conserved sequences from sequences that have been conserved between *C. elegans* and *C. briggsae* by chance (3). The addition of the *C. remanei* and *C. sp. PB2801* genomes should also help to locate transcription factor binding sites within CNSs.

5. Detecting Syntenic Blocks and Chromosomal Rearrangements

5.1. Creating Whole Genome Alignments

Comparative studies of genome structure and evolution, and of conserved protein-coding and noncoding regions, typically rely on high-quality nucleotide-level alignments. Most pairwise genome alignments begin with a genome-wide set of local alignments—which are then stitched together into non-overlapping blocks that can be used for an analysis of synteny.

Among the many alignment algorithms available, the WABA (42) and BLASTZ (32) algorithms have been particularly effective at aligning the *C. elegans* and *C. briggsae* genomes (2). Unlike algorithms tuned for aligning protein-coding regions, WABA and BLASTZ can effectively align regions of genomes that are evolving neutrally. WABA generates a set of local alignments between genomes, and discriminates between coding and noncoding regions, and between “strong” and “weak” alignments. As such, WABA is useful for identifying conserved coding regions between genomes. WABA alignments between *C. elegans* and *C. briggsae* cover 65% of their genomes (2). Surprisingly, alignments between *C. elegans* and *C. briggsae* are as common in introns and intergenic regions as in coding regions, so are useful for locating potential regulatory sequences (31). WABA alignments between *C. elegans* and *C. briggsae* can be downloaded from WormBase (<http://www.wormbase.org>) (4).

The BLASTZ algorithm (32) is derived from gapped BLAST (21), with enhancements for aligning long sequences. In initial tests, WABA and BLASTZ performed comparably at aligning the *C. elegans* and *C. briggsae* genomes, resulting in 65 and 56% non-overlapping coverage, respectively. Generating whole genome alignments between species using BLASTZ or WABA requires considerable computational resources and time. Nonetheless, alignments can be generated between a small set of regions and a target genome quickly and easily on modest hardware. For example, upstream regions of *C. elegans* genes could be rapidly aligned to a newly sequenced genome to search for candidate regulatory sequences.

New methods have recently been developed to align multiple genomes simultaneously, such as MULTIZ (43). MULTIZ uses pairwise alignments between two genomes to guide subsequent alignments with other genomes. MULTIZ offers several advantages over pairwise alignments, including independence from the reference genome (*C. elegans* in the case of *Caenorhabditis* species). Pairwise and multiple alignments between the *C. elegans*, *C. briggsae*, *C. remanei*, *C. sp.* PB2801, and *C. japonica* genomes will be made available from WormBase (4) in the near future.

5.2. Detecting Syntenic Blocks

Synteny is the colinearity of genes between species. For whole genome analyses, this definition is often extended to include colinearity of segments of chro-

mosomes, not just of those regions containing genes. Analyses of synteny typically rely on nucleotide-level alignments or on sets of anchor points between two genomes. Nucleotide-level alignments usually require some postprocessing in order to generate a set of syntenic blocks. This is because a region in one genome may have multiple alignments that map to several different regions in the second genome. Furthermore, an alignment may be much shorter than the syntenic block that contains it—if the alignment algorithm halted prematurely in a poorly conserved region. No hard and fast rules exist for deriving syntenic blocks from nucleotide-level alignments; the process usually needs to be modified according to the degree of similarity between genomes.

The *C. briggsae*–*C. elegans* synteny analysis illustrates one method of postprocessing a set of whole genome nucleotide-level alignments to find larger syntenic blocks (2). First, adjacent alignments that had the same WABA status (which can be “weak,” “strong,” or “coding”) were merged (Fig. 3). Second, *C. elegans* blocks that contained nucleotide-level alignments to more than five *C. briggsae* blocks, and *C. briggsae* blocks that contained alignments to more than five *C. elegans* blocks, were discarded. Third, a “simple merge” algorithm was used to merge adjacent blocks that had conserved order in *C. elegans* and *C. briggsae*. Finally, a dynamic programming algorithm was used to find the longest series of blocks having conserved order in *C. elegans* and *C. briggsae*. This algorithm bypassed short nonsyntenic blocks, enabling the creation of longer syntenic blocks. For each *C. briggsae* supercontig, the algorithm first found the longest series of contiguous blocks and merged these, and then found the next longest series using the blocks left over from the first iteration. This continued until no blocks remained. During this process, the merging of blocks that had conserved order was restricted such that the resultant merged blocks had to have similar sizes in the two species. This was done by restricting nonsyntenic gaps to less than 100 kb, and by not allowing gaps that would cause a greater than fivefold expansion of a syntenic block in either genome.

After careful merging of nucleotide-level alignments into syntenic blocks, it may be necessary to discard very small blocks. After merging *C. briggsae*–*C. elegans* WABA alignments into syntenic blocks, there was a large spike in the distribution of block size at approx 1250 bp. Many blocks of this size contained a single nucleotide-level alignment that correlated poorly with the positions of known orthologous genes. We considered these blocks to be unreliable, and so excluded all blocks of less than 1850 bp from the final analysis. This reduced the syntenic coverage of the genome by only 1.5%, but excluded 64% of merged blocks made by the “simple merge” step.

The true extent of synteny is underestimated when unfinished genomes are compared. This is because the ends of contigs are necessarily considered to be synteny breakpoints—when they may in fact be part of a contiguous syntenic

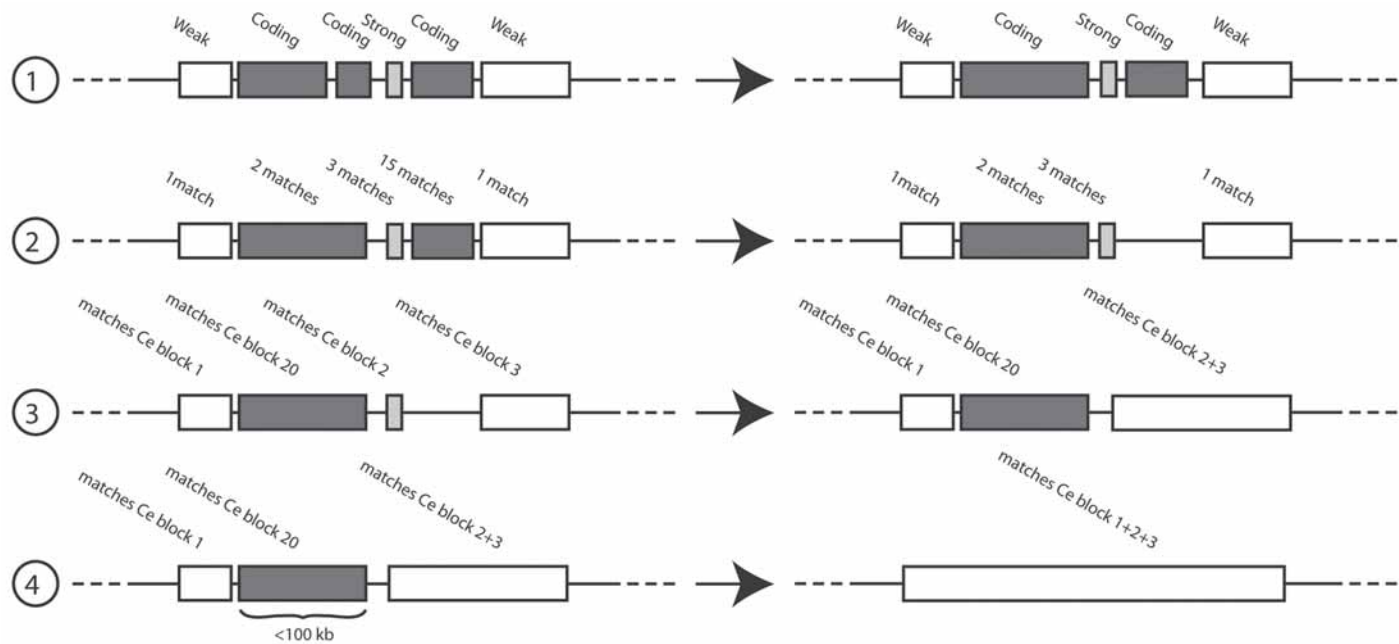


Fig. 3. The approach used by Stein et al. (2) to derive *Caenorhabditis elegans*–*Caenorhabditis briggsae* syntenic blocks from genome-wide nucleotide-level alignments. First, adjacent alignments that had the same WABA status (“weak,” “strong,” or “coding”) were merged. Second, blocks that matched more than five blocks in the other species were discarded. Third, a “simple merge” algorithm was used to merge adjacent blocks that had conserved order in *C. elegans* and *C. briggsae*. Finally, a dynamic programming algorithm was used to find the longest series of blocks having conserved order in *C. elegans* and *C. briggsae*. This algorithm bypassed short nonsyntenic blocks of less than 100 kb, enabling the creation of longer syntenic blocks.

block. Comparisons to multiple species can resolve these ambiguities, because a contiguous breakpoint will rarely occur in the same relative position in several species' genome assemblies.

Analysis of synteny does not require whole genome nucleotide-level alignments. Synteny can be analyzed using any set of anchor points between two genomes, such as the positions of orthologous genes. As part of the *C. briggsae*–*C. elegans* analysis, a companion method to the alignment-based synteny analysis was developed, which used a conservative set of orthologous gene assignments (2). This algorithm approaches the sensitivity of synteny analysis based on nucleotide-level alignments—but is faster and easier to implement. Orthologous genes (or other anchor features) are numbered in order along the contigs or chromosomes. Syntenic blocks are then defined by a series of orthologous genes that have conserved order between the two genomes. This approach overlooks syntenic regions that extend beyond gene boundaries, such as upstream regulatory sequences at the end of a syntenic block.

Newly sequenced (and less studied genomes) often lack contig-to-chromosome mappings. When comparing such a genome assembly to a finished genome that does have chromosome mappings, one can analyze synteny by numbering anchor points (nucleotide-level alignments or genes) in the newly sequenced genome according to the position of their orthologs in the second (finished) genome. This allows each contig of the newly sequenced genome to be placed and oriented on the second genome, according to the longest series of anchor points that have conserved order within that contig. This strategy works well for genomes with many anchor points and a large average contig size, but accuracy drops dramatically as the average contig size decreases.

5.3. Characterizing Breaks in Colinearity

With a set of merged non-overlapping syntenic blocks, it is relatively straightforward to classify breaks in colinearity. Each end of a syntenic block represents a potential break in synteny. By examining the neighboring blocks of each segment in relation to a reference genome, each block can be classified as an inversion, a transposition, or a reciprocal translocation (2). Given an arrangement of syntenic blocks in *C. briggsae*:

===== **a/b**—————**c/d** =====

then the block **bc** was classified as an inversion if **a** was adjacent to **c** in *C. elegans*, and **b** was adjacent to **d** in *C. elegans*. On the other hand, the *C. briggsae* block **bc** was classified as a transposition if **a** and **d** were adjacent in *C. elegans*, and a reciprocal event could not be identified in *C. elegans*. Alter-

natively, the block **bc** was classified as involving one or two reciprocal translocations if another *C. briggsae* breakpoint was found:

===== e/f —————

and **a** was adjacent to **f** in *C. elegans*, or **e** was adjacent to **d** in *C. elegans*.

5.4. Visualizing Syntenic Blocks

WormBase (<http://www.wormbase.org> [4]) provides several options for viewing syntenic blocks between *C. elegans* and *C. briggsae* graphically. First, in the Genome Browser, syntenic blocks are displayed under the “briggsae alignments” track. Within the syntenic blocks, WABA-classified “coding” regions appear in dark blue, “strong” regions in light blue, and “weak” regions in gray. Because syntenic blocks can contain gaps for which there is no WABA nucleotide-level alignment, they may not be fully continuous. A dashed gray line represents these gaps in a syntenic block. Syntenic blocks can be viewed from either the perspective of the *C. elegans* or the *C. briggsae* genome. Clicking on a syntenic block takes the user to the Synteny Browser. The Synteny Browser displays the relationship of syntenic blocks to both the *C. elegans* and *C. briggsae* genomes simultaneously.

6. Conclusion

The initial pairwise comparisons between *C. elegans* and *C. briggsae* built a foundation for the steps necessary for whole genome analyses of nematode genomes. We will soon have a data set of five different *Caenorhabditis* genomes. Comparison of these five genomes will both require and stimulate the development of novel techniques that make use of multispecies comparisons to predict genes and their regulatory sequences, to define orthologs and gene families, and to identify regions of chromosomal synteny. These analyses will provide the opportunity to study the evolution of nematode genomes in unprecedented detail.

Acknowledgments

Avril Coghlan is supported by the Wellcome Trust. Jason Stajich is supported by a National Science Foundation predoctoral fellowship. Todd Harris is supported by National Human Genome Research Institute grant P41-HG02223 to the WormBase Consortium.

References

1. Kiontke, K. and Sudhaus, W. (2006) Ecology of *Caenorhabditis* species, WormBook, ed. The *C. elegans* Research Community WormBook, doi/10.1895/wormbook.1.37.1. Available at <http://www.wormbook.org>.

- 1a. The *C. elegans* Sequencing Consortium (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012–2018.
2. Stein, L. D., Bao, Z., Blasiar, D., et al. (2003) The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. *PLoS Biol.* **1**, E45.
3. Sternberg, P. W., Waterston, R. H., Speith, J., Eddy, S. R., and Wilson, R. K. (2003) Genome sequence of additional *Caenorhabditis* species: enhancing the utility of *C. elegans* as a model organism. National Human Genome Research Institute. <http://www.genome.gov/Pages/Research/Sequencing/SeqProposals/CaenorhabditisSEQ.pdf> Accessed: April 2006.
4. Chen, N., Harris, T. W., Antoshechkin, I., et al. (2005) WormBase: a comprehensive data resource for *Caenorhabditis* biology and genomics. *Nucleic Acids Res.* **33**, D383–D389.
5. Worm ENCODE Consortium (2005) Worm ENCODE White Paper. http://www.wormbase.org/support_files/2005/worm_encode_white_paper.txt. Accessed: April 2006.
6. Salamov, A. A. and Solovyev, V. V. (2000) *Ab initio* gene finding in *Drosophila* genomic DNA. *Genome Res.* **10**, 516–522.
7. Howe, K. L., Chothia, T., and Durbin, R. (2002) GAZE: a generic framework for the integration of gene-prediction data by dynamic programming. *Genome Res.* **12**, 1418–1427.
8. Wei, C., Lamesch, P., Arumugam, M., et al. (2005) Closing in on the *C. elegans* ORFeome by cloning TWINSCAN predictions. *Genome Res.* **15**, 577–582.
9. Brent, M. R. and Guigó, R. (2004) Recent advances in gene structure prediction. *Curr. Opin. Struct. Biol.* **14**, 264–272.
10. Korf, I., Flicek, P., Duan, D., and Brent, M. R. (2001) Integrating genomic homology into gene structure prediction. *Bioinformatics* **17 Suppl 1**, S140–S148.
11. Siepel, A. C. and Haussler, D. (2004) Computational identification of evolutionarily conserved exons. In: *RECOMB 2004: Proceedings of the Eighth Annual International Conference on Molecular Biology*, March 27–31, 2004, San Diego, CA, ACM Press, New York, NY, pp. 177–186.
12. Pedersen, J. S. and Hein, J. (2003) Gene finding with a hidden Markov model of genome structure and evolution. *Bioinformatics* **19**, 219–227.
13. Liu, J., Tzou, P., Hill, R. J., and Sternberg, P. W. (1999) Structural requirements for the tissue-specific and tissue-general functions of the *Caenorhabditis elegans* epidermal growth factor LIN-3. *Genetics* **153**, 1257–1269.
14. Cho, S., Jin, S. W., Cohen, A., and Ellis, R. E. (2004) A phylogeny of *Caenorhabditis* reveals frequent loss of introns during nematode evolution. *Genome Res.* **14**, 1207–1220.
15. Wistrand, M. and Sonnhammer, E. L. (2005) Improved profile HMM performance by assessment of critical algorithmic features in SAM and HMMER. *BMC Bioinformatics* **6**, 99.
16. Long, M. (2001) Evolution of novel genes. *Curr. Opin. Genet. Dev.* **11**, 673–680.
17. Domazet-Lošo, T. and Tautz, D. (2003) An evolutionary analysis of orphan genes in *Drosophila*. *Genome Res.* **13**, 2213–2219.

18. Coghlan, A. and Wolfe, K. H. (2004) Origins of recently gained introns in *Caenorhabditis*. *Proc. Natl. Acad. Sci. USA* **101**, 11,362–11,367.
19. Kiontke, K., Gavin, N. P., Raynes, Y., Roehrig, C., Piano, F., and Fitch, D. H. (2004) *Caenorhabditis* phylogeny predicts convergence of hermaphroditism and extensive intron loss. *Proc. Natl. Acad. Sci. USA* **101**, 9003–9008.
20. Llopart, A., Comeron, J. M., Brunet, F.G., Lachaise, D., and Long, M. (2002) Intron presence-absence polymorphism in *Drosophila* driven by positive Darwinian selection. *Proc. Natl. Acad. Sci. USA* **99**, 8121–8126.
21. Altschul, S. F., Madden, T. L., Schäffer, A. A., et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402.
22. Remm, M., Storm, C. E., and Sonnhammer, E. L. (2001) Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J. Mol. Biol.* **314**, 1041–1052.
23. Li, L., Stoeckert, C. J., Jr., and Roos, D. S. (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* **13**, 2178–2189.
24. van Dongen, S. (2000) Ph.D. thesis. *Graph clustering by flow simulation*. Dutch National Research Institute for Mathematics and Computer Science, University of Utrecht.
25. Eisen, J. A. (1998) Phylogenomics: improving functional predictions for uncharacterized genes by evolutionary analysis. *Genome Res.* **8**, 163–167.
26. Ohno, S. (1970) *Evolution by Gene Duplication*. Springer Verlag, NY.
27. Lespinet, O., Wolf, Y. I., Koonin, E. V., and Aravind, L. (2002) The role of lineage-specific gene family expansion in the evolution of eukaryotes. *Genome Res.* **12**, 1048–1059.
28. Chen, N., Pai, S., Zhao, Z., et al. (2005) Identification of a nematode chemosensory gene family. *Proc. Natl. Acad. Sci. USA* **102**, 146–151.
29. Enright, A. J., Van Dongen, S., and Ouzounis, C. A. (2002) An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* **30**, 1575–1584.
30. Ghedin, E., Wang, S., Foster, J. M., and Slatko, B. E. (2004) First sequenced genome of a parasitic nematode. *Trends Parasitol.* **20**, 151–153.
31. Webb, C. T., Shabalina, S. A., Ogurtsov, A. Y., and Kondrashov, A. S. (2002) Analysis of similarity within 142 pairs of orthologous intergenic regions of *Caenorhabditis elegans* and *Caenorhabditis briggsae*. *Nucleic Acids Res.* **30**, 1233–1239.
32. Schwartz, S., Kent, W. J., Smit, A., et al. (2003) Human-mouse alignments with BLASTZ. *Genome Res.* **13**, 103–107.
33. Batzoglou, S., Pachter, L., Mesirov, J. P., Berger, B., and Lander, E. S. (2000) Human and mouse gene structure: comparative analysis and application to exon prediction. *Genome Res.* **10**, 950–958.
34. Gaudet, J., Muttumu, S., Horner, M., and Mango, S. E. (2004) Whole-genome analysis of temporal gene expression during foregut development. *PLoS Biol.* **2**, E352.

35. Kirouac, M. and Sternberg, P. W. (2003) *cis*-Regulatory control of three cell fate-specific genes in vulval organogenesis of *Caenorhabditis elegans* and *C. briggsae*. *Dev. Biol.* **257**, 85–103.
36. Matys, V., Fricke, E., Geffers, R., et al. (2003) TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res.* **31**, 374–378.
37. GuhaThakurta, D., Schriefer, L. A., Waterston, R. H., and Stormo, G. D. (2004) Novel transcription regulatory elements in *Caenorhabditis elegans* muscle genes. *Genome Res.* **14**, 2457–2468.
38. Ao, W., Gaudet, J., Kent, W. J., Muttumu, S., and Mango, S.E. (2004) Environmentally induced foregut remodeling by PHA-4/FoxA and DAF-12/NHR. *Science* **305**, 1743–1746.
39. Tompa, M., Li, N., Bailey, T. L., et al. (2005) Assessing computational tools for the discovery of transcription factor binding sites. *Nat. Biotechnol.* **23**, 137–144.
40. Pavesi, G., Mereghetti, P., Mauri, G., and Pesole, G. (2004) Weeder Web: discovery of transcription factor binding sites in a set of sequences from co-regulated genes. *Nucleic Acids Res.* **32**, W199–W203.
41. Wang, X., Greenberg, J. F., and Chamberlin, H. M. (2004) Evolution of regulatory elements producing a conserved gene expression pattern in *Caenorhabditis*. *Evol. Dev.* **6**, 237–245.
42. 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.* **10**, 1115–1125.
43. Blanchette, M., Kent, W. J., Riemer, C., et al. (2004) Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.* **14**, 708–715.

C. elegans

METHODS IN MOLECULAR BIOLOGY™

John M. Walker, SERIES EDITOR

384. **Capillary Electrophoresis: Methods and Protocols**, edited by *Philippe Schmitt-Kopplin*, 2007
383. **Cancer Genomics and Proteomics: Methods and Protocols**, edited by *Paul B. Fisher*, 2007
382. **Microarrays, Second Edition: Volume 2, Applications and Data Analysis**, edited by *Jang B. Rampil*, 2007
381. **Microarrays, Second Edition: Volume 1, Synthesis Methods**, edited by *Jang B. Rampil*, 2007
380. **Immunological Tolerance: Methods and Protocols**, edited by *Paul J. Fairchild*, 2007
379. **Glycovirolgy Protocols**, edited by *Richard J. Sugrue*, 2007
378. **Monoclonal Antibodies: Methods and Protocols**, edited by *Maher Albitar*, 2007
377. **Microarray Data Analysis: Methods and Applications**, edited by *Michael J. Korenberg*, 2007
376. **Linkage Disequilibrium and Association Mapping: Analysis and Application**, edited by *Andrew R. Collins*, 2007
375. **In Vitro Transcription and Translation Protocols: Second Edition**, edited by *Guido Grandi*, 2007
374. **Quantum Dots: Methods and Protocols**, edited by *Charles Z. Hotz and Marcel Bruchez*, 2007
373. **Pyrosequencing® Protocols**, edited by *Sharon Marsh*, 2007
372. **Mitochondrial Genomics and Proteomics Protocols**, edited by *Dario Leister and Johannes Herrmann*, 2007
371. **Biological Aging: Methods and Protocols**, edited by *Trygve O. Tollefsbol*, 2007
370. **Adhesion Protein Protocols, Second Edition**, edited by *Amanda S. Coutts*, 2007
369. **Electron Microscopy: Methods and Protocols, Second Edition**, edited by *John Kuo*, 2007
368. **Cryopreservation and Freeze-Drying Protocols, Second Edition**, edited by *John G. Day and Glyn Stacey*, 2007
367. **Mass Spectrometry Data Analysis in Proteomics**, edited by *Rune Matthiesen*, 2007
366. **Cardiac Gene Expression: Methods and Protocols**, edited by *Jun Zhang and Gregg Rokosh*, 2007
365. **Protein Phosphatase Protocols**: edited by *Greg Moorhead*, 2007
364. **Macromolecular Crystallography Protocols: Volume 2, Structure Determination**, edited by *Sylvie Doublé*, 2007
363. **Macromolecular Crystallography Protocols: Volume 1, Preparation and Crystallization of Macromolecules**, edited by *Sylvie Doublé*, 2007
362. **Circadian Rhythms: Methods and Protocols**, edited by *Ezio Rosato*, 2007
361. **Target Discovery and Validation Reviews and Protocols: Emerging Molecular Targets and Treatment Options, Volume 2**, edited by *Mouldy Sioud*, 2007
360. **Target Discovery and Validation Reviews and Protocols: Emerging Strategies for Targets and Biomarker Discovery, Volume 1**, edited by *Mouldy Sioud*, 2007
359. **Quantitative Proteomics by Mass Spectrometry**, edited by *Salvatore Sechi*, 2007
358. **Metabolomics: Methods and Protocols**, edited by *Wolfram Weckwerth*, 2007
357. **Cardiovascular Proteomics: Methods and Protocols**, edited by *Fernando Vivanco*, 2006
356. **High-Content Screening: A Powerful Approach to Systems Cell Biology and Drug Discovery**, edited by *D. Lansing Taylor, Jeffrey Haskins, and Ken Guiliano*, 2007
355. **Plant Proteomics: Methods and Protocols**, edited by *Hervé Thiellement, Michel Zivy, Catherine Damerval, and Valerie Mechin*, 2006
354. **Plant-Pathogen Interactions: Methods and Protocols**, edited by *Pamela C. Ronald*, 2006
353. **DNA Analysis by Nonradioactive Probes: Methods and Protocols**, edited by *Elena Hilario and John. F. MacKay*, 2006
352. **Protein Engineering Protocols**, edited by *Kristian Müller and Katja Arndt*, 2006
351. **C. elegans: Methods and Applications**, edited by *Kevin Strange*, 2006
350. **Protein Folding Protocols**, edited by *Yawen Bai and Ruth Nussinov* 2007
349. **YAC Protocols, Second Edition**, edited by *Alasdair MacKenzie*, 2006
348. **Nuclear Transfer Protocols: Cell Reprogramming and Transgenesis**, edited by *Paul J. Verma and Alan Trounson*, 2006
347. **Glycobiology Protocols**, edited by *Inka Brockhausen*, 2006
346. **Dictyostelium discoideum Protocols**, edited by *Ludwig Eichinger and Francisco Rivero*, 2006
345. **Diagnostic Bacteriology Protocols, Second Edition**, edited by *Louise O'Connor*, 2006
344. **Agrobacterium Protocols, Second Edition: Volume 2**, edited by *Kan Wang*, 2006
343. **Agrobacterium Protocols, Second Edition: Volume 1**, edited by *Kan Wang*, 2006
342. **MicroRNA Protocols**, edited by *Shao-Yao Ying*, 2006
341. **Cell-Cell Interactions: Methods and Protocols**, edited by *Sean P. Colgan*, 2006
340. **Protein Design: Methods and Applications**, edited by *Raphael Guerois and Manuela López de la Paz*, 2006
339. **Microchip Capillary Electrophoresis: Methods and Protocols**, edited by *Charles S. Henry*, 2006
338. **Gene Mapping, Discovery, and Expression: Methods and Protocols**, edited by *M. Bina*, 2006
337. **Ion Channels: Methods and Protocols**, edited by *James D. Stockand and Mark S. Shapiro*, 2006
336. **Clinical Applications of PCR, Second Edition**, edited by *Y. M. Dennis Lo, Rossa W. K. Chiu, and K. C. Allen Chan*, 2006
335. **Fluorescent Energy Transfer Nucleic Acid Probes: Designs and Protocols**, edited by *Vladimir V. Didenko*, 2006

METHODS IN MOLECULAR BIOLOGY™

C. elegans

Methods and Applications

Edited by

Kevin Strange

*Departments of Anesthesiology, Molecular Physiology
and Biophysics, and Pharmacology
Vanderbilt University Medical Center
Nashville, TN*


HUMANA PRESS  TOTOWA, NEW JERSEY

© 2006 Humana Press Inc.
999 Riverview Drive, Suite 208
Totowa, New Jersey 07512

humanapress.com

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording, or otherwise without written permission from the Publisher. Methods in Molecular Biology™ is a trademark of The Humana Press Inc.

All papers, comments, opinions, conclusions, or recommendations are those of the author(s), and do not necessarily reflect the views of the publisher.

This publication is printed on acid-free paper. 
ANSI Z39.48-1984 (American Standards Institute)

Permanence of Paper for Printed Library Materials.

Cover illustration: Fig. 2, Chapter 15, "Preservation of *C. elegans* Tissue Via High-Pressure Freezing and Freeze-Substitution for Ultrastructural Analysis and Immunocytochemistry," by Robby M. Weimer.

Production Editor: Amy Thau

Cover design by Patricia F. Cleary.

For additional copies, pricing for bulk purchases, and/or information about other Humana titles, contact Humana at the above address or at any of the following numbers: Tel.: 973-256-1699; Fax: 973-256-8341; E-mail: orders@humanapress.com; or visit our Website: www.humanapress.com

Photocopy Authorization Policy:

Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by Humana Press Inc., provided that the base fee of US \$30.00 per copy is paid directly to the Copyright Clearance Center at 222 Rosewood Drive, Danvers, MA 01923. For those organizations that have been granted a photocopy license from the CCC, a separate system of payment has been arranged and is acceptable to Humana Press Inc. The fee code for users of the Transactional Reporting Service is: [1-58829-597-4/06 \$30.00].

Printed in the United States of America. 10 9 8 7 6 5 4 3 2 1

eISBN: 1-59745-151-7

ISSN: 1064-3745

Library of Congress Cataloging in Publication Data

C. elegans : methods and applications / edited by Kevin Strange.

p. ; cm. -- (Methods in molecular biology ; 351)

Includes bibliographical references and index.

ISBN 1-58829-597-4 (alk. paper)

1. *Caenorhabditis elegans*. 2. *Caenorhabditis elegans*--Genetics.

3. Molecular biology. I. Strange, Kevin, 1955- . II. Methods in molecular biology (Clifton, N.J.) ; v. 351.

[DNLM: 1. *Caenorhabditis elegans*--physiology. 2. *Caenorhabditis elegans*--ultrastructure. W1 ME9616J v.351 2006 / QX 203

C1112 2006]

QL391.N4E64 2006

572.8'1257--dc22

2006000939

Preface

Molecular biology has driven a powerful reductionist, or “molecule-centric,” approach to biological research in the last half of the 20th century. Reductionism is the attempt to explain complex phenomena by defining the functional properties of the individual components of the system. Bloom (1) has referred to the post-genome sequencing era as the end of “naïve reductionism.” Reductionist methods will continue to be an essential element of all biological research efforts, but “naïve reductionism,” the belief that reductionism alone can lead to a complete understanding of living organisms, is not tenable. Organisms are clearly much more than the sum of their parts, and the behavior of complex physiological processes cannot be understood simply by knowing how the parts work in isolation.

Systems biology has emerged in the wake of genome sequencing as the successor to reductionism (2–5). The “systems” of systems biology are defined over a wide span of complexity ranging from two macromolecules that interact to carry out a specific task to whole organisms. Systems biology is integrative and seeks to understand and predict the behavior or “emergent” properties of complex, multicomponent biological processes. A systems-level characterization of a biological process addresses the following three main questions: (1) What are the parts of the system (i.e., the genes and the proteins they encode)? (2) How do the parts work? (3) How do the parts work together to accomplish a task?

Nonmammalian model organisms such as *Escherichia coli*, *Saccharomyces*, *Caenorhabditis elegans*, *Drosophila*, *Danio rerio*, and the plant *Arabidopsis* have become cornerstones of systems biology research. They have been likened to the Rosetta stone (4), which provided modern scholars with the tools needed to decipher Egyptian hieroglyphics. Similarly, model organisms provide investigators the experimental tools necessary to decipher the genetic code that underlies complex physiological processes common to all life.

C. elegans provides a particularly striking example of the experimental utility of nonmammalian model organisms. Worms have a short life cycle (2–3 d at 25°C), produce large numbers of offspring by sexual reproduction, and can be cultured easily and inexpensively in the laboratory. Sexual reproduction occurs by self-fertilization in hermaphrodites or by mating with males. Self-fertilization allows homozygous animals to breed true and greatly facilitates the isolation and maintenance of mutant strains, whereas mating with males allows mutations to be moved between strains. The reproductive and laboratory culture characteristics

of *C. elegans* make it an exceptionally powerful model system for forward genetic analysis. Mutagenesis and genetic screening allow unbiased identification of genes underlying a biological process of interest, allow the genes to be ordered into pathways, and can provide important and novel mechanistic insights into the molecular structure and function of proteins.

In addition to forward genetic tractability, *C. elegans* also has a fully sequenced and well-annotated genome. Genomic sequence and virtually all other biological data on this organism are assembled in readily accessible public databases (e.g., WormBase; <http://www.wormbase.org>). Numerous reagents, including mutant worm strains and cosmid and YAC clones spanning the genome, are freely available through public resources. Creation of transgenic worms is relatively easy, inexpensive, and rapid requiring little more than injection of transgenes into the animal's gonad or bombardment with DNA-coated microparticles. *C. elegans* gene expression can be specifically and potently targeted for knockdown using RNA interference, either at the single worm level by injection of double-stranded RNA, or at the population level by feeding worms double-stranded RNA-producing bacteria. Finally, *C. elegans* is a highly differentiated animal but is comprised of less than 1000 somatic cells. This relatively simple anatomy greatly facilitates the study of biological processes and has made it possible to trace the lineage of every adult cell beginning with the first cell division (6,7), and to generate a complete wiring diagram of the 302 neuron adult hermaphrodite nervous system (8).

A wealth of methodology for the study of *C. elegans* is described online and in the printed literature. The goal of *C. elegans: Methods and Applications* is to provide overviews and detailed step-by-step descriptions of newer and state-of-the-art methods utilized in the field. These include tools essential for forward and reverse genetic analysis, data mining and comparative genomics strategies, electron and fluorescence microscopy methods, automated imaging methods for worm behavioral analysis, functional genomics strategies, and methods for physiological analyses including somatic cell culture, toxicity assays, electrophysiology, and in vivo imaging of intracellular Ca^{2+} and pH using genetically encoded fluorescent indicator proteins. It is my hope that this book will be of use to both experts and newcomers to the field, not only as a step-by-step guide, but also as a roadmap to show what is possible with *C. elegans* and what has yet to be discovered.

Kevin Strange

References

1. Bloom, F. E. (2001) What does it all mean to you? *J. Neurosci.* **21**, 8304–8305.
2. Kitano, H. (2002) Looking beyond the details: a rise in system-oriented approaches in genetics and molecular biology. *Curr. Genet.* **41**, 1–10.
3. Kitano, H. (2002) Systems biology: a brief overview. *Science* **295**, 1662–1664.
4. Ideker, T., Galitski, T., and Hood, L. (2001) A new approach to decoding life: systems biology. *Annu. Rev. Genomics Hum. Genet.* **2**, 343–372.
5. Pennisi, E. (2003) Systems biology. Tracing life's circuitry. *Science* **302**, 1646–1649.
6. Sulston, J. E. and Horvitz, H. R. (1977) Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev. Biol.* **56**, 110–156.
7. Sulston, J. E., Schierenberg, E., White, J. G., and Thomson, J. N. (1983) The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64–119.
8. White, J. G., Southgate, E., Thomson, J. N., and Brenner, S. (1986) The structure of the nervous system of the nematode *C. elegans*. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **314**, 1–340.

Contents

Preface	v
Contributors	xi
1 An Overview of <i>C. elegans</i> Biology	1
Kevin Strange	
2 Comparative Genomics in <i>C. elegans</i> , <i>C. briggsae</i> , and Other <i>Caenorhabditis</i> Species	13
Avril Coghlan, Jason E. Stajich, and Todd W. Harris	
3 WormBase: <i>Methods for Data Mining</i> and <i>Comparative Genomics</i>	31
Todd W. Harris and Lincoln D. Stein	
4 <i>C. elegans</i> Deletion Mutant Screening	51
Robert J. Barstead and Donald G. Moerman	
5 Insertional Mutagenesis in <i>C. elegans</i> Using the <i>Drosophila</i> Transposon <i>Mos1</i> : <i>A Method for the Rapid Identification</i> of Mutated Genes	59
Jean-Louis Bessereau	
6 Single-Nucleotide Polymorphism Mapping	75
M. Wayne Davis and Marc Hammarlund	
7 Creation of Transgenic Lines Using Microparticle Bombardment Methods	93
Vida Praitis	
8 Construction of Plasmids for RNA Interference and In Vitro Transcription of Double-Stranded RNA	109
Lisa Timmons	
9 Delivery Methods for RNA Interference in <i>C. elegans</i>	119
Lisa Timmons	
10 Functional Genomic Approaches in <i>C. elegans</i>	127
Todd Lamitina	
11 Assays for Toxicity Studies in <i>C. elegans</i> With Bt Crystal Proteins	139
Larry J. Bischof, Danielle L. Huffman, and Raffi V. Aroian	
12 Fluorescent Reporter Methods	155
Harald Hutter	

13	Electrophysiological Analysis of Neuronal and Muscle Function in <i>C. elegans</i>	175
	Michael M. Francis and Andres Villu Maricq	
14	Sperm and Oocyte Isolation Methods for Biochemical and Proteomic Analysis	193
	Michael A. Miller	
15	Preservation of <i>C. elegans</i> Tissue Via High-Pressure Freezing and Freeze-Substitution for Ultrastructural Analysis and Immunocytochemistry	203
	Robby M. Weimer	
16	Intracellular pH Measurements In Vivo Using Green Fluorescent Protein Variants	223
	Keith Nehrke	
17	Automated Imaging of <i>C. elegans</i> Behavior	241
	Christopher J. Cronin, Zhaoyang Feng, and William R. Schafer	
18	Intracellular Ca ²⁺ Imaging in <i>C. elegans</i>	253
	Rex A. Kerr and William R. Schafer	
19	In Vitro Culture of <i>C. elegans</i> Somatic Cells	265
	Kevin Strange and Rebecca Morrison	
20	Techniques for Analysis, Sorting, and Dispensing of <i>C. elegans</i> on the COPAS™ Flow-Sorting System	275
	Rock Pulak	
	Index	287

Contributors

- RAFFI V. AROIAN • *Section of Cell and Developmental Biology, University of California, San Diego, La Jolla, CA*
- ROBERT J. BARSTEAD • *Oklahoma Medical Research Foundation, Oklahoma City, OK*
- JEAN-LOUIS BESSEREAU • *Department of Biology, Ecole Normale Supérieure, Paris, France*
- LARRY J. BISCHOF • *Section of Cell and Developmental Biology, University of California, San Diego, La Jolla, CA*
- AVRIL COGHLAN • *Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Belfield, Ireland; and Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK*
- CHRISTOPHER J. CRONIN • *HHMI and Division of Biology, California Institute of Technology, Pasadena, CA*
- M. WAYNE DAVIS • *Department of Biology, University of Utah, Salt Lake City, UT*
- ZHAOYANG FENG • *Life Science Institute, University of Michigan, Ann Arbor, MI*
- MICHAEL M. FRANCIS • *Department of Biology, University of Utah, Salt Lake City, UT*
- MARC HAMMARLUND • *Department of Biology, University of Utah, Salt Lake City, UT*
- TODD W. HARRIS • *Cold Spring Harbor Laboratory, Cold Spring Harbor, NY*
- DANIELLE L. HUFFMAN • *Section of Cell and Developmental Biology, University of California, San Diego, La Jolla, CA*
- HARALD HUTTER • *Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia, Canada*
- REX A. KERR • *Division of Biological Sciences, University of California, San Diego, La Jolla, CA*
- TODD LAMITINA • *Department of Anesthesiology, Vanderbilt University Medical Center, Nashville, TN*
- ANDRES VILLU MARICQ • *Department of Biology, University of Utah, Salt Lake City, UT*
- MICHAEL A. MILLER • *Department of Cell Biology, University of Alabama at Birmingham, Birmingham, AL*

- DONALD G. MOERMAN • *Genome Sciences Centre, BC Cancer Agency, Vancouver, BC, Canada*
- REBECCA MORRISON • *Department of Anesthesiology, Vanderbilt University Medical Center, Nashville, TN*
- KEITH NEHRKE • *Nephrology Unit, Department of Medicine and Department of Pharmacology and Physiology, University of Rochester Medical Center, Rochester, NY*
- VIDA PRAITIS • *Department of Biology, Grinnell College, Grinnell, IA*
- ROCK PULAK • *Life Sciences Technology Group, Union Biometrica Inc., Holliston, MA*
- WILLIAM R. SCHAFER • *Division of Biological Sciences, University of California, San Diego, La Jolla, CA*
- JASON E. STAJICH • *Department of Molecular Genetics and Microbiology, Duke University, Durham, NC*
- LINCOLN D. STEIN • *Cold Spring Harbor Laboratory, Cold Spring Harbor, NY*
- KEVIN STRANGE • *Departments of Anesthesiology, Molecular Physiology and Biophysics, and Pharmacology, Vanderbilt University Medical Center, Nashville, TN*
- LISA TIMMONS • *Department of Molecular Biosciences, University of Kansas, Lawrence, KS*
- ROBBY M. WEIMER • *Cold Spring Harbor Laboratory, Cold Spring Harbor, NY*

Index

A

- AceDB, *see* WormBase
- Acetylcholine receptor, neuromuscular junction analysis, 184
- Automated tracking,
 - Converter for file conversion, 247, 248
 - Lineup, 248, 249
 - locomotion mutants, 241, 242
 - machine vision system,
 - calibration, 245, 246
 - hardware, 243
 - preparation, 244
 - recording, 246, 247
 - starting, 244
 - track switch, 246
 - materials, 242, 243
 - Miner for behavior analysis, 249, 250

B

- Bacillus thuringiensis* crystal proteins,
 - Caenorhabditis elegans* toxicity assays, advantages in, 139, 140
 - brood size assay, 147, 148, 152
 - Escherichia coli*-expressed Cry5B plate assay, 149
 - larva growth assay in L1 stage, imaging, 145, 146, 150
 - measurement and data analysis, 146, 147
 - setup, 145, 150
 - LC₅₀ assay, 148, 149, 152
 - materials, 142–144
 - principles, 141, 142
 - resistant mutants, 140
 - worm bleaching and preparation, 144
 - toxicity mechanisms, 140
- Behavior, *see* Automated tracking
- Biolistics, *see* Transgenic worms
- BLASTZ, whole genome sequence alignment, 22

C

- Caenorhabditis elegans*,
 - anatomy,
 - digestive tract, 6
 - gonad, 6, 7
 - gross morphology, 4
 - kidney, 5, 6
 - muscle, 4, 5
 - nervous system, 5
 - skin, 4
 - behavior, *see* Automated tracking
 - calcium imaging, *see* Calcium imaging
 - comparative genomics, *see* Comparative genomics, *Caenorhabditis*
 - culture, *see* Culture, *Caenorhabditis elegans*
 - databases and resources, overview, 7–9
 - WormBase, *see* WormBase
 - electrophysiology, *see* Electrophysiology, *Caenorhabditis elegans*
 - flow sorting, *see* COPAS™ Biosort instrument
 - gene mapping, *see* Single nucleotide polymorphism mapping
 - immunocytochemistry, *see* Immunoelectron microscopy
 - intracellular pH measurement, *see* pHluorin
 - life cycle, 2, 3
 - model organism establishment, 1, 2
 - mutagenesis, *see* Deletion mutant screening; Insertional mutagenesis
 - RNA interference, *see* RNA interference
 - toxicity assays, *see* *Bacillus thuringiensis* crystal proteins
 - transgenesis, *see* Transgenic worms
- Calcium imaging,
 - applications in *Caenorhabditis elegans*, 253, 254
 - cameleons, 254
 - camgaroo, 254
 - data analysis, 258–263

- fluorescence microscopy, 257, 258, 260, 261
 - indicator types, 254, 255
 - materials, 255, 256, 259, 260
 - pericam, 254, 255
 - sample preparation, 256, 257, 260
 - Chromosome mapping, *see* Single nucleotide polymorphism mapping
 - Comparative genomics, *Caenorhabditis*,
 - annotation of related genes, 15
 - gene family identification, 17–19
 - gene prediction in related genomes, 13–15
 - gene structure comparison in related genomes, 15, 16
 - ortholog identification, 16, 17
 - paralog identification, 16, 17
 - regulatory element prediction, 19, 21
 - syntenic blocks,
 - colinearity break characterization, 25, 26
 - detection, 22, 23, 25
 - visualization, 26
 - whole genome sequence alignment, 22
 - WormBase tools, 45, 46
 - Confocal microscopy, *see* Fluorescence microscopy
 - COPAS™ Biosort instrument,
 - data output and archiving, 278
 - flow cytometer comparison, 277
 - instrumentation, 279–281
 - principles, 275, 276
 - rare event sorting, 284, 285
 - software, 277
 - sorting parameters, 276, 277
 - stress response analysis, 281–283
 - Cry proteins, *see* *Bacillus thuringiensis* crystal proteins
 - Culture, *Caenorhabditis elegans*,
 - cell culture,
 - dissociated embryo cell preparation, 270–272
 - egg isolation, 270
 - materials, 268
 - peanut lectin-coated culture vesicle preparation, 269, 272
 - rationale and applications, 265, 266, 268
 - worm culture synchronization, 269
 - deletion mutant screening, 52
 - DNA microarray and large-scale culture, 129, 130, 136
 - microparticle bombardment culture, 99, 100, 104
 - overview, 3
- D**
- Deletion mutant screening,
 - applications, 51
 - culture of worms, 52
 - deletion stabilization, 56–58
 - DNA template detection, 53, 54
 - library culture, 53, 57
 - materials, 52
 - mutagenesis, 52, 53, 57
 - overview, 51, 52
 - polymerase chain reaction detection of
 - deletions, 54, 55
 - sibling selection, 55, 56
 - DNA microarray,
 - applications, 127
 - culture of worms, 129, 130, 136
 - experimental design, 129, 136
 - materials, 128
 - poly(A) RNA preparation, 131
 - principles, 127, 128
 - RNA purification, 136
 - Double-stranded RNA, *see* RNA interference
- E**
- Electron microscopy, *see* Immuno-electron microscopy
 - Electrophysiology, *Caenorhabditis elegans*,
 - advantages of study, 175, 176
 - electropharyngeogram recording, 176
 - limitations, 187, 188
 - neuromuscular junction,
 - acetylcholine receptor analysis, 184
 - features, 180–182
 - neurotransmitter release analysis, 184, 185
 - synaptic current measurement, 182–184
 - neuron studies,
 - AVA interneuron, 185, 186
 - primary sensory neuron recordings, 186
 - patch-clamp,
 - cultured cell studies, 186, 187
 - principles, 176, 177
 - recording, 179, 180
 - slit-worm dissection, 177, 178
 - prospects for study, 187, 188
 - Embryo, dissociated embryo cell preparation and culture, 270–272
 - Evogene, gene prediction in related genomes, 15

F

- Fgenesh, gene prediction in related genomes, 14
- Flow sorting, *see* COPAS™ Biosorter
- Fluorescence microscopy,
 - calcium imaging, *see* Calcium imaging
 - genome-wide phenotypic screening, *see* RNA interference
 - intracellular pH measurement, *see* pHluorin
 - live worm imaging of dyes and fluorescent proteins,
 - anesthesia, 160
 - confocal microscopy, 164
 - DiI staining of neurons, 162
 - eggshell puncturing with laser, 161, 162
 - materials, 158, 159
 - mounting,
 - agar pad preparation, 159, 160
 - individual worms, 160
 - multiple worms, 160
 - recovery from slides, 161
 - time-lapse recording, 164, 165
 - wide-field fluorescence microscopy, 162, 163
- Fluorescence polarization-template directed incorporation (FP-TDI), single nucleotide polymorphism detection, 87
- Fluorescence resonance energy transfer (FRET), cameleons, 254
- Fluorescent dyes,
 - anatomical structure labeling, 157
 - applications, 224
 - imaging,
 - anesthesia, 160
 - confocal microscopy, 164
 - DiI staining of neurons, 162
 - eggshell puncturing with laser, 161, 162
 - materials, 158, 159
 - mounting,
 - agar pad preparation, 159, 160
 - individual worms, 160
 - multiple worms, 160
 - recovery from slides, 161
 - time-lapse recording, 164, 165
 - wide-field fluorescence microscopy, 162, 163
 - intracellular pH, 225

- FP-TDI, *see* Fluorescence polarization-template directed incorporation
- Freeze substitution, *see* Immuno-electron microscopy
- FRET, *see* Fluorescence resonance energy transfer

G

- Gametes, *see* Oocyte; Sperm
- Gene gun, *see* Transgenic worms
- Gene mapping, *see* Single nucleotide polymorphism mapping
- GeneFinder, gene prediction in related genomes, 14
- Genome microarray analysis, *see* DNA microarray
- GFP, *see* Green fluorescent protein
- Gonad,
 - anatomy, 6, 7
 - gametes, *see* Oocyte; Sperm
- Gonad, liberation, 198, 199
- Green fluorescent protein (GFP),
 - calcium imaging, *see* Calcium imaging
 - cotransformation marker, 95
 - dye limitations in structure labeling, 157
 - genome-wide phenotypic screening, *see* RNA interference
 - imaging,
 - anesthesia, 160
 - confocal microscopy, 164
 - materials, 158, 159
 - mounting,
 - agar pad preparation, 159, 160
 - individual worms, 160
 - multiple worms, 160
 - recovery from slides, 161
 - time-lapse recording, 164, 165
 - wide-field fluorescence microscopy, 162, 163
 - intracellular pH measurement, *see* pHluorin
 - reporter applications, 156
 - variants, 156
 - vectors, 157, 158

H

- High-pressure freezing, *see* Immuno-electron microscopy

I

- Immuno-electron microscopy,
 challenges of *Caenorhabditis elegans*, 203
 freeze substitution,
 apparatus set-up, 210, 211, 216, 217
 immunocytochemistry samples,
 211–213, 217, 218
 materials, 207, 208, 215
 morphometric analysis samples, 211,
 217, 218
 principles, 204
 high-pressure freezing,
 freezing, 209
 liquid nitrogen transfer of sample, 209
 loading worms into freeze chambers,
 208, 209, 216
 materials, 206, 207, 215
 principles, 204
 postembedding immunogold staining,
 205, 208, 213–216, 218

- InDel detection, single nucleotide
 polymorphism detection, 87, 88

- Insertional mutagenesis,
Caenorhabditis elegans advantages, 59
 transposons,
 insertion localization with inverse
 polymerase chain reaction,
 67–69, 71
 materials, 61, 62, 69, 70
Mos1 features, 60
 mutagenesis, 62, 63, 70, 71
 mutagenic insert identification,
 65–67, 71
 overview, 60, 61
 transposition rate measurement,
 63–65, 71

L

- Locomotion, *see* Automated tracking

M

- Machine vision, *see* Automated tracking
 Microparticle bombardment, *see* Transgenic
 worms
Mos1, *see* Insertional mutagenesis
 MULTIZ, whole genome sequence
 alignment, 22
 Mutagenesis, *see* Deletion mutant
 screening; Insertional mutagenesis

N

- Neuromuscular junction, *see*
 Electrophysiology,
Caenorhabditis elegans

O

- Oocyte,
 gonad liberation, 198, 199
 isolation,
 centrifugation, 198
 dissection, 197, 200
 filtration, 197
 materials, 194, 195, 199
 storage, 198
 worm culture, 197, 199
 proteome analysis advantages, 194
 Ortholog, identification in related genomes,
 16, 17

P

- Paralog, identification in related genomes,
 16, 17
 Patch-clamp, *see* Electrophysiology,
Caenorhabditis elegans
 PCR, *see* Polymerase chain reaction
 pHluorin,
 applications, 225
 intracellular pH measurements in vivo,
 agarose pad,
 preparation, 228, 229
 worm mounting, 229
 dual excitation ratio imaging,
 acquisition parameters, 232, 233
 calibration *in situ*, 235
 data acquisition and analysis, 233,
 234
 instrumentation, 229, 230, 232
 vendors, 231
 expression plasmids, 226–228, 235,
 236
 materials, 225, 226
 Polymerase chain reaction (PCR),
 deletion detection, 54, 55
 single nucleotide polymorphism
 genotyping,
 chromosome mapping, 79, 80, 90
 interval mapping, 81, 91
 primer design, 83, 84
 transposon insertion localization with
 inverse polymerase chain reaction,
 67–69, 71

Postembedding immunogold staining, *see*
Immuno-electron microscopy

R

Ratiometric fluorescence, *see* Calcium
imaging; pHLuorin

Regulatory element, computational
prediction, 19, 21

Restricted fragment length polymorphism
(RFLP), single nucleotide
polymorphism detection, 87

RFLP, *see* Restricted fragment length
polymorphism

RNA interference (RNAi),

discovery, 109

double-stranded RNA delivery,

feeding, 123–125

injection, 121, 122, 124

materials, 120, 121, 124

overview of approaches, 110, 119, 120

soaking, 122–124

transcription in vivo, 124

genome-wide phenotypic screening,
applications, 127

bacteria feeding library,

availability, 132

culture, 134

expansion from microtiter plates
to Omnitrays, 133, 134, 137

replication and storage, 132, 133,
136

culture of worms, 135, 137

materials, 128

phenotypic assays, 135, 137

principles, 127, 128

screening, 136

plasmid construction,

DNA templates, 111–114

double-stranded RNA production by in
vitro transcription, 115–117

materials, 111–114

target selection, 111

RNAi, *see* RNA interference

S

Single nucleotide polymorphism (SNP)
mapping,

advantages, 76

applications, 75–77

challenges,

CB4856 background effects, 89

complex genotypes, 88–90

overview, 76

chromosome mapping,

data analysis, 80

mating, 78, 79, 90

polymerase chain reaction, 79, 80, 90

primers, 79

detection techniques,

fluorescence polarization-template

directed incorporation, 87

InDel detection, 87, 88

restricted fragment length

polymorphism, 87

interval mapping,

data analysis, 82, 83

directed mapping,

CB4856 background, 86

overview, 84, 85

three-factor mapping, 85, 86

two-factor mapping, 85

mating, 80, 81, 91

polymerase chain reaction, 81, 91

primer design, 83, 84

materials, 77, 78, 89, 90

SNP mapping, *see* Single nucleotide

polymorphism mapping

Sperm,

isolation,

centrifugation, 196, 199

filtration, 196, 199

materials, 194, 195, 199

sex separation, 196

sperm collection, 196, 199

storage, 197

worm culture, 195, 199

proteome analysis advantages, 194

Syntenic blocks,

colinearity break characterization, 25, 26

detection, 22, 23, 25

visualization, 26

T

Tracking, *see* Automated tracking

Transgenic worms,

germ-line injection approach, 93, 94

- microparticle bombardment,
 - advantages and limitations, 95
 - applications, 94
 - approaches, 97
 - Bio-Rad Helios Gene Gun,
 - bead preparation, 103, 104
 - postbombardment care of worms, 103
 - transformation, 103, 104
 - Biolistic PDS100/He Particle delivery System,
 - bead preparation, 100, 101, 104
 - with Hepta adapter, 101, 104
 - without Hepta adapter, 101, 104
 - postbombardment care of worms, 101, 102, 104
 - cotransformation markers, 95–97
 - DNA preparation, 99
 - materials, 97–99
 - Schnabel alternative microparticle bombardment device,
 - bead preparation, 102, 104
 - postbombardment care of worms, 102–104
 - transformation, 102, 104
 - worm culture, 99, 100, 104
 - Transposon insertional mutagenesis, *see* Insertional mutagenesis
 - TRIBE-MCL, gene family identification in related genomes, 17
 - TWINSKAN, gene prediction in related genomes, 14, 15
- W**
- WABA, whole genome sequence alignment, 22
 - WormBase,
 - access options, 32
 - AceDB data model, 33, 34
 - citation, 33
 - classes, 34, 36
 - comparative genomics tools, 45, 46
 - content overview, 32
 - data freezes, 33
 - Genome Browser display of remote annotations, 45
 - hardware requirements, 37
 - linking, 44
 - object identifiers, 34
 - origins, 31
 - oversight, 31, 32
 - programmatic data mining,
 - AceDB programmatic access using Ace Perl, 41, 42, 46, 47
 - GFF programmatic access using Bio::DB::GFF, 42, 43, 46–48
 - Web interface, 39–41, 46, 47
 - query languages,
 - AceDB, 38
 - WormBase, 37
 - versions and release cycles, 32, 33
 - Web-based data-mining options, 37

C. elegans

Methods and Applications

Edited by

Kevin Strange

*Departments of Anesthesiology, Molecular Physiology and Biophysics, and Pharmacology,
Vanderbilt University Medical Center, Nashville, TN*

Nonmammalian model organisms have become a cornerstone of systems biology research. Like the Rosetta Stone, which enabled modern scholars to decode ancient Egyptian hieroglyphics, model organisms enable biologists to decipher the genetic code underlying the complex physiological processes common to all life. *C. elegans* provides a particularly striking example of the experimental utility of model organisms. Genetic, molecular and systems biological characterization of this organism is greatly facilitated by its short life cycle and high rate reproduction, and by the ease with which it can be cultured in the laboratory. *C. elegans* also has a fully sequenced and well-annotated genome, which is assembled in readily accessible public databases along with virtually all other biological data on this organism.

C. elegans: Methods and Applications aims to enhance the readily available methodologies for the study of *C. elegans* by providing overviews and concise, step-by-step descriptions of the most state-of-the-art methods currently being utilized in the field. Some of the methods presented include forward and reverse genetic analysis, data mining and comparative genomics strategies, electron and fluorescence microscopy methods, automated imaging methods for worm behavioral analysis, functional genomics strategies, and cutting-edge methods for physiological analyses. This comprehensive collection of methods, written by experts in the field, endeavors to serve as a roadmap for researchers, illustrating the type of research that is possible with *C. elegans*, and illuminating its potential as a vehicle for future discovery.

FEATURES

- Presents methods for light and electron microscopy
- Examines quantitative and brightfield fluorescence imaging
- Considers RNAi methods specific to *C. elegans*, both at the single worm level and at the population level
- Explores methods for physiological analyses, including somatic cell culture, toxicity assays, electrophysiology and in vivo imaging of intracellular Ca²⁺ and pH

CONTENTS

An Overview of *C. elegans* Biology. Comparative Genomics in *C. elegans*, *C. briggsae*, and Other *Caenorhabditis* Species. Worm-Base: Methods for Data Mining and Comparative Genomics. *C. elegans* Deletion Mutant Screening. Insertional Mutagenesis in *C. elegans* Using the *Drosophila* Transposon *Mos1*: A Method for the Rapid Identification of Mutated Genes. Single-Nucleotide Polymorphism Mapping. Creation of Transgenic Lines Using Micro-particle Bombardment Methods. Construction of Plasmids for RNA Interference and In Vitro Transcription of Double-Stranded RNA. Delivery Methods for RNA Interference in *C. elegans*. Functional Genomic Approaches in *C. elegans*. Assays for Toxicity Studies in *C. elegans* With Bt Crystal Proteins. Fluorescent Reporter

Methods. Electrophysiological Analysis of Neuronal and Muscle Function in *C. elegans*. Sperm and Oocyte Isolation Methods for Biochemical and Proteomic Analysis. Preservation of *C. elegans* Tissue Via High-Pressure Freezing and Freeze-Substitution for Ultrastructural Analysis and Immunocytochemistry. Intracellular pH Measurements In Vivo Using Green Fluorescent Protein Variants. Automated Imaging of *C. elegans* Behavior. Intracellular Ca²⁺ Imaging in *C. elegans*. In Vitro Culture of *C. elegans* Somatic Cells. Techniques for Analysis, Sorting, and Dispensing of *C. elegans* on the COPAS™ Flow-Sorting System. Index.

ISBN 1-58829-597-4



9 0000

