invited review

Physiol Genomics 13: 15–24, 2003;
10.1152/physiolgenomics.00075.2002.

Super models

Maureen M. Barr
School of Pharmacy, University of Wisconsin-Madison, Madison, Wisconsin 53705

Barr, Maureen M. Super models. *Physiol Genomics* 13: 15–24, 2003; 10.1152/physiolgenomics.00075.2002.—Model organisms have been used over a century to understand basic, conserved biological processes. The study of these experimental systems began with genetics and development, moved into molecular and cellular biology, and most recently propelled into functional genomics and proteomics. The goal of this review is simple: to discuss the place of model organisms in "The Age of the Ome": the genome, the transcriptome, and the proteome. This review will address the following questions. What exactly is a model organism? What characteristics make an excellent model system? Using the yeast *Saccharomyces cerevisiae* and the nematode *Caenorhabditis elegans* as examples, this review will discuss these issues with the aim of demonstrating how model organisms remain indispensable scientific tools for understanding complex biological pathways and human disease.

Saccharomyces cerevisiae; *Caenorhabditis elegans*; genetics; genomics; proteomics

The world's mine oyster, which I with sword will open—
William Shakespeare (1564–1616)

METAPHORS AND SIMILES are useful literary devices for describing and comprehending our world. Model organisms have been used over a century to understand basic, conserved biological processes. The study of these experimental systems began with genetics and development, moved into molecular and cellular biology, and most recently propelled into functional genomics and proteomics. The goal of this review is simple, to discuss the place of model organisms in "The Age of the Ome": the genome, the transcriptome, and the proteome. I hope to address the following questions. What exactly is a model organism? What characteristics make an excellent model system? Using the yeast *Saccharomyces cerevisiae* and the nematode *Caenorhabditis elegans* as examples, in this review I will discuss these issues with the aim of demonstrating how model organisms remain indispensable scientific tools for understanding complex biological pathways and human disease.

Complete genomic sequence provides an endless source of information for understanding the molecular make up of an organism. Sequencing projects have provided the scientific community with genetic codes ranging from bacteria to human. Complete, or nearly complete, genome sequence is known for an elite group of eukaryotic organisms, including three vertebrates: the yeasts *S. cerevisiae* (37) and *Schizosaccharomyces pombe* (101), the nematode *C. elegans* (12), the fruit fly *Drosophila melanogaster* (1), the human malaria parasite *Plasmodium falciparum* and mosquito carrier *Anopheles gambiae* (33, 45), the mustard cress *Arabidopsis thaliana* (3), domestic rice *Oryza sativa* (36, 105), the puffer fish *Fugu rubripes* (2), the mouse *Mus musculus* (97), and the human *Homo sapiens* (55, 95). The knowledge of full genome sequence information is drastically changing experimental approaches and is rapidly shaping the future of scientific research. Sequence databases provide a starting point for data mining of genomic information, and there is a wealth of Internet resources available to link DNA sequence information with the study of model organisms (Table 1).

The number of predicted human genes is estimated to be between 26,000 and 40,000 (55, 95), although this number is controversial (23) and considered to be an underestimate by some groups (24, 44, 102). Analysis of the mouse genome indicates a similar number (97). The genomes of the budding yeast *S. cerevisiae* and fission yeast *S. pombe* encode about 6,400 and 4,900 products, respectively (21). *C. elegans* and *Drosophila* boast 20,000 and 14,000 predicted genes (21). Are we really only two times more complicated than a worm? Unlikely. A single gene may encode multiple proteins, for example, by alternative splicing of its mRNA transcript or by alternative start or stop sites. The proteins encoded by our genome are often more complex, possessing multiple functional domains. Protein modifications (such as phosphorylation or glycosylation) and subcellular localization profoundly determine function. Understanding the function of an organism's DNA (genome), RNA (transcriptome), and protein (proteome) components requires a holistic approach. The study of a model organism, sometimes accused to be reductionist, is well suited for addressing this daunting task.

Article published online before print. See web site for date of publication (http://physiolgenomics.physiology.org).

Address for reprint requests and other correspondence: M. Barr, School of Pharmacy, Univ. of Wisconsin-Madison, 777 Highland Ave., Madison, WI 53705 (E-mail: mmbarr@pharmacy.wisc.edu).

Table 1. *Online resources*

Saccharomyces Genome Database	http://genome-www.stanford.edu/Saccharomyces/
The Zebrafish Information Network	http://zfin.org/ http://www.sanger.ac.uk/Projects/D_rerio/
Flybase (*Drosophila*)	http://www.flybase.org/
Wormbase (*C. elegans*)	http://www.wormbase.org/
C. elegans WWW server	http://elegans.swmed.edu
Mouse Genome Informatics	http://www.informatics.jax.org/
The *S. pombe* Genome Sequencing Project	http://www.sanger.ac.uk/Projects/S_pombe/
The *Arabidopsis* Information Resource	http://www.arabidopsis.org/
Xenbase: A *Xenopus* web resource	http://xenbase.org/
S. cerevisiae, S. pombe, and *C. elegans* (databases now require a license)	https://www.incyte.com/proteome/databases.jsp

What Is A Model Organism?

Origins: before genome vs. after genome. A model organism may have its experimental origins in one of two time periods: before or after the conception of the Human Genome Project (before genome, BG, or after genome, AG). BG models were developed to study classic and molecular genetics, development, and/or physiology. For example, the study of inheritance began in *Drosophila* in 1910 with T. H. Morgan's laboratory discovering a spontaneous mutant with white eye color. The classic eukaryotic BG models are *S. cerevisiae* (and "the other yeast," *S. pombe*), *C. elegans*, *D. melanogaster*, *Danio rerio*, and *M. musculus*. For brevity, this review will highlight *S. cerevisiae* and *C. elegans*. BG models are also fundamental tools in the AG world. The sequencing projects of bacteria, yeast, and, especially, the worm have provided the framework for other genome sequencing projects, including the Human Genome Project. The newly completed genome of the mouse, *M. musculus*, the mammalian model organism of choice for medical and behavioral research, will help decipher the human genome. The most useful model organisms are instrumental in both BG and AG scientific inquiry.

AG models are being selected largely based on their potential to contribute to improvements in human health research and to advance biomedical and industrial progress. An AG model may not be a useful experimental system in and of itself. To interpret the human genome, Sydney Brenner (also known for his role as founding father of the *C. elegans* field) initiated sequencing the genome of the puffer fish *F. rubripes* as a compact model vertebrate genome (2, 11; for review, see Ref. 94). The genome sequencing and analysis of this AG model is informative in the study of DNA sequence features and comparative genomics. On the other hand, *Fugu* does not shed additional light on the biological functions of, and complex interactions between, genes. Nevertheless, the genomes of AG organisms are worth sequencing to fill in gaps in the evolutionary tree. Vertebrate genome sequencing and interspecies comparison is essential for comparative and evolutionary genomics.

The genomes of hundreds of microorganisms have been sequenced for human health and commercial purposes. The three major infectious diseases in humans (malaria, tuberculosis, and AIDS) are caused by intracellular pathogens whose entire genome sequence is available [*P. falciparum* (33), *Mycobacterium tuberculosis* (20, 30), and human immunodeficiency virus (67), respectively]. Coupling genome information from the malaria parasite, vector, and host (the malaria parasite *P. falciparum*, the vector mosquito *A. gambiae*, and the human host, respectively) will allow for rational drug design and discovery as well as the development of insecticides and effective vaccines. While it is not feasible to employ this tripartite malaria system in the laboratory, *Drosophila* may be useful model in genetic studies of malaria parasite (75) and in comparative genomics for *A. gambiae* (18, 71, 73).

Model organism characteristics: survival of the fittest meets symbiosis. With an estimated one million nematode species (8), why is the free-living roundworm *C. elegans* commonly known in the scientific community as "The Worm"? To be an exemplary model organism, one must meet rigorous criteria yet possess unique traits (Table 2). First and foremost, it should be amenable to both forward (phenotype to gene) and reverse (gene to phenotype) genetic approaches (Fig. 1). Classic forward genetic characteristics include the ability to determine dominance, complementation, and recombination. Reverse genetic tools to inactivate genes must be available. Other prerequisites for model organism consideration include a sequenced, or soon to be sequenced, genome and the ability to generate transgenic animals (or cells, in the case of yeast) by DNA transformation. These basic tools are essential to dissect and understand gene function.

A model organism must also be utilitarian. Practical qualities include a short generation time, small size, and both ease and reasonable cost of maintenance. For

Table 2. *Universal features of a model organism*

Genetically Amenable
Tools for forward and reverse genetics
Transgenic Capability
DNA transformation
Genome sequence
Practicalities: short generation time, small size, cost of maintenance
Unique properties
Critical research mass

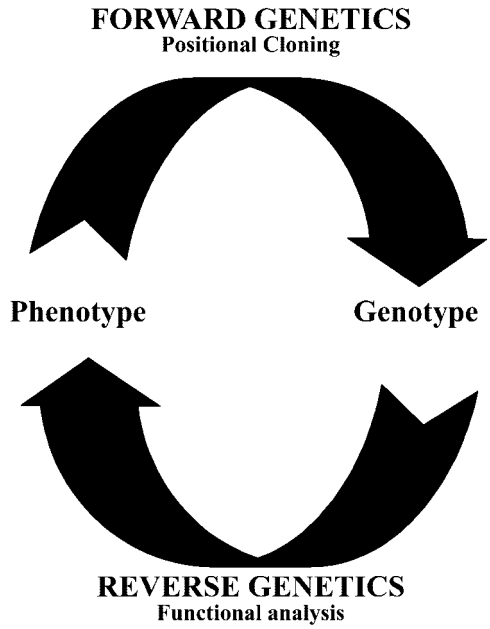

Fig. 1. Flow of information in reverse and forward genetics approaches. Forward genetics refers to identifying genes based on mutant phenotype followed by positional cloning of mutated genes (phenotype to genotype). Reverse genetics refers to the functional analysis of a gene of known molecular identity (genotype to phenotype).

example, *C. elegans* has a 3-day generation time at 20°C, is 1 mm long, is easily maintained, and requires only *Escherichia coli* for sustenance (Table 3). The frog *Xenopus laevis* is a superb animal to study embryonic vertebrate development (98), yet its painfully slow 3-yr generation time rules out genetic analysis. For this reason, its precocious cousin *X. tropicalis* with a relatively shorter generation time (5 mo) is being developed as a complementary system to study developmental genetics (60a).

Other practicalities include a critical mass of researchers and strong resources (Table 3). *Drosophila* is the most popular model organism, with about 1,500 fly laboratories. *S. cerevisiae* is the most cited, with over 50,000 references listed in PubMed (**http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed**). Resources available to the *S. cerevisiae* and *C. elegans* community are astounding: both have federally funded stock centers and genome-wide knockout projects, and *S. cerevisiae* has a genome-wide yeast two-hybrid project. A great challenge is compiling data from both individual laboratories and large research projects. The model organism community has responded by establishing online databases "providing the research community with accurate, current, accessible information concerning the genetics, genomics and biology" of yeast, worms, flies, fish, and mice (Table 1; the preceding quote is from **http://www.wormbase.org/about/about_WormBase.html**).

Finally, an ideal model organism must possess unique characteristics that simplify analysis of the biological problem of interest. These special strengths often enable researchers to address fundamental questions in a new light and in a manner not possible in other systems. *S. cerevisiae* boasts the awesome power of yeast genetics. *Drosophila* claims sophisticated genetics, polytene chromosomes, and a wealth of information on developmental biology. A rapid generation time, limited number of cells, and hermaphroditism distinguish *C. elegans*. External fertilization and transparency makes the zebrafish a powerful vertebrate model. The mouse is a small animal for which to model the basis of mammalian development and human disease.

All model organisms have significant weaknesses as well (Table 3). One major weakness of all genetically amenable model organisms is that for which they were selected: speed of development. Clearly all organisms do not share this trait, as exemplified by ourselves! The simplicity of yeast is also a drawback: unicellularity is not conducive to study complex developmental phenomena. Similarly, *C. elegans* does not have many specialized tissues. Both *C. elegans* and *D. melanogaster* are highly specialized organisms with genes that are often very divergent at the sequence level from mammalian homologs. Relative to the other nonmammalian BG models, *D. rerio* is in its infancy, and this is reflected in the short supply of available technologies and resources. Taking into account both the strengths and weakness of a specific animal illustrates why it is so critical to use multiple model systems and perhaps invest in developing genetic and genomic resources for models on the fringe and in exploring new models. Historically, it is clear that integrating the information gained from several model organisms has tremendous power.

The Yeasts S. cerevisiae and S. pombe

The budding yeast *S. cerevisiae* and the fission yeast *S. pombe* are single-celled fungi with distinct lifestyles and evolutionarily diverged genomes. However, they share many powerful molecular genetic tools and have a rapid generation time (doubling approximately every few hours), making them amenable to both classic genetics (31) and high-throughput genomic approaches (53). *S. cerevisiae* was the first eukaryote to be transformed by plasmids (6), to have targeted gene disruption via homologous recombination (74), and to have a completely sequenced genome (37).

As with any completed genome, the major challenge is to dissect the function of the 6,400 and 4,900 genes in *S. cerevisiae* and *S. pombe*, respectively, 30% of which have completely unknown function (Table 4). How is function ascertained? There are a few basic approaches used in any model organism but most advanced in *S. cerevisiae*: turn the gene off (knockout) or on (overexpression), determine gene expression pattern and protein subcellular localization, identify interacting proteins, and analyze enzymatic function. Genomics is concerned with the former three and proteomics, the latter. The endurance and relevance of model organisms in the Ome Age is perhaps best exemplified by the

Table 3. *Nonmammalian model organisms*

	S. cerevisiae	*C. elegans*	*D. melanogaster*	*D. rerio*
Cellular organization	unicellular	multicellular	complex multicellular	complex multicellular; vertebrate
Ploidy	haploid/diploid	diploid	diploid	diploid
Generation time	hours (doubling time)	3 days (20°C)	12 days (25°C)	3 mo
Maintainance	simple, inexpensive	simple, inexpensive	simple, inexpensive	requires space for tanks
Transgenics	yes	yes	yes	yes
Gene inactivation	Homologous recombination	RNAi, transposon insertion, chemical mutagen	RNAi, transposon insertion, chemical and X-ray mutagen	morpholinos, RNAi, insertional mutagenesis
Mutants	many	many	many	yes
Cell culture	yes	yes (17)	yes	yes
Long-term storage	indefinite (stabs and glycerol stocks)	months (starved as dauers); indefinite (−80°C)	no, requires constant passaging	sperm
Resources*	genome-wide knockout, genome-wide 2H, stock center	genome-wide RNAi in progress, Genome-wide knockout in progress Stock center, EM resource	stock centers	stock center in progress
Strengths	unicellular, homologous recombination, powerful genomic and proteomic technologies	invariant cell lineages, small number of cells, neuronal connectivity known, transparent	sophisticated genetics, well-characterized development, battery of mutants, 2-component inducible system, homologous recombination	external fertilization, high fecundity, transparent, vertebrate
Weaknesses	unicellular	no homologous recombination	No embryo freezing	no homologous recombination, tetrapolid
Community (no. labs)	624[1]	259[2]	1,481[3]	308[4]

*With the exception of *D. rerio,* the genome sequence of these model organisms is complete or nearly complete. [1]http://genome-www.stanford.edu/Saccharomyces. [2]http://elegans.swmed.edu/Worm_labs/. [3]http://www.flybase.org/. [4]http://zfin.org/ZFIN/. Modified from the table of D. Valle et al. at http://www.nih.gov/science/models/nmm/appc1.html, with permission.

prosperous marriage of genetics, genomics, and proteomics in *S. cerevisiae*.

Genetics. Classic or forward genetics moves from function (as determined by mutation) to gene identification. Conversely, both reverse genetics and high-throughput genomics progresses from gene to function. The 2001 Nobel Prize in Physiology or Medicine was awarded to three biologists who defined the molecular basis of cell division. Leland Hartwell identified cell-division-cycle (CDC) genes in *S. cerevisiae* (42) while Paul Nurse isolated similar genes in *S. pombe* (61). Later, Nurse demonstrated evolutionary conservation of cell cycle regulators by identifying a human CDK (cyclin-dependent kinase). Using biochemical approaches in sea urchin and *Xenopus* oocytes, Timothy Hunt discovered cyclins, proteins that regulate CDK function. This pioneering work done in the 1970s and 1980s gave rise to and paved the way for the cell cycle field (62). Amazingly, the draft human genome sequence (55, 95) provided only a few new cyclins and no new CDKs, indicating that classic experimentation in yeast had identified nearly all cell cycle regulators (60).

Genomics and proteomics: the yeast paradigm for high-throughput approaches. To determine the function of all the yeast gene products, the *Saccharomyces* Genome Deletion Project has deleted ~6,200 genes whose open reading frames (ORFs) encode at least 100 amino acids from start to stop (100). Each deletion

Table 4. *Genome size*

Organism	Genome Size, Mb (euchromatin + heterochromatin)	No. Genes	No. Genes with Known/Inferred Function
S. cerevisiae (Ref. 21)	12	6,419	5,260
S. pombe (Ref. 21)	13.8 (Ref. 101)	4,959	3,699
C. elegans (Ref. 21)	97	20,317	12,484
D. melanogaster	180 (120 + 60) (BDGP)	13,601 (BDGP)	~3,000 (NCBI)
D. rerio	1,700		
M. musculus	3,088 (NCBI)	30,000 predicted	Less than 5,000
H. sapiens	3,000	26,000–40,000	

BDGP, Berkeley *Drosophila* Genome Project; NCBI, National Center for Biotechnology Information.

strain is generated via PCR targeting and labeled with a unique "bar code" that enables simultaneous analysis of many deletion mutants (77). This collection is available for distribution (**http://www-sequence.stanford.edu/group/yeast_deletion_project/**). Approximately 1,100 genes are essential, ~5,100 haploid null mutants are viable, and about 1/3 of genes possess unknown function (Table 4). To explore genetic redundancy and genetic pathways, synthetic genetic array (SGA) analysis was developed to systematically construct and analyze double mutants (91). Synthetic lethality/sickness of double mutants (with neither single mutant exhibiting a lethal phenotype) revealed functional relationships between 204 genes. The SGA approach may be potentially applied to higher eukaryotes where high-throughput gene knockout technologies are possible.

Genome annotation methods predict gene number, accuracy, and function. Gene annotation is not direct, and gene identification hinges on computer predictions, homology to other organisms, expression, and classic techniques of gene cloning and random cDNA analysis. Depending on the criteria, gene numbers may be overpredicted or underestimated. Individual ORFs may have incorrect intron/exon predictions, may not encode functional genes, or may be overlooked. In theory, gene prediction in *S. cerevisiae*, an organism with a compact genome, few introns, and rare trans-splicing, should be straightforward, yet this is not the case. For example, Kumar et al. (52) surveyed ~40% of the *S. cerevisiae* genome and identified 137 overlooked genes in *S. cerevisiae* using a combination of random transposon insertion and gene trapping, microarray-based expression analysis, and genome-wide homology comparisons. Comparison of the *S. cerevisiae* genome with sequence data from other fungi will be useful in homology searches (63, 106). From this lesson learned from yeast, it is clear that reevaluating and reannotating a genome using a variety of complementary techniques is a necessity.

Gene expression may be used to correlate or infer function. *S. cerevisiae* DNA microarrays have been used to investigate differential gene expression in various cellular events, including metabolism (25), cell cycle (16, 79), meiosis (19, 68), transcriptional regulation (9), DNA damage (7, 49), and DNA replication (48, 70). In a genomic/proteomic hybrid approach, DNA microarrays have also been employed for study of protein-DNA interactions. For example, transcription-factor and origin of replication binding sites have been identified using this approach (48, 103). DNA microarrays are useful in any organism, and technological advances made using yeast will assist similar studies of the human genome.

Protein-protein interactions have been studied via two high-throughput detection methods: yeast two-hybrid systems (47, 90, 93) and protein complex purification using mass spectrometry (34, 43). The yeast two-hybrid system detects protein-protein interactions via transcriptional activation of a reporter gene by two fusion proteins: one fused to a DNA binding domain and the other to an activation domain (28). The two-hybrid system has been used in individual assays and in systematic arrays. Although arrays have been exploited to look at protein family interactions of yeast, *C. elegans*, and *Drosophila*, large-scale, genome-wide two-hybrid arrays have only been performed in yeast (for review, see Ref. 92). Large-scale studies of protein complexes, typically composed of three or more components, have also been executed in yeast (for review, see Ref. 54). As with any high-throughput method, demonstration of physiological relevance of these interactions is strengthened by independent and complementary methods (96) and ultimately proceeds more slowly in a case-by-case basis.

Protein localization within a cell often correlates with and indicates function. Subcellular localization is typically determined by a reporter such as lacZ or green fluorescent protein (GFP) or epitope tag. Kumar et al. (51) performed a yeast proteome-wide analysis of protein localization using a combination of directed epitope-tagging of PCR amplified ORFs and random tagging by transposon mutagenesis. Subcellular localization of 2,744 proteins, of which 1,000 had previously unknown function, was determined and cataloged in a searchable database available at the Yale Genome Analysis Center home page at **http://ygac.med.yale.edu**.

Biochemical protein activity has also been systematically accessed in the *S. cerevisiae* proteome (108). 5,800 ORFs were cloned and overexpressed. The resulting tagged proteins were purified, printed onto microarrays, and assayed for ability to interact with a protein (calmodulin) or phospholipids. In theory, these proteome chips can be screened for interactions with individual proteins, drugs, and lipids as well as for enzymatic function.

The Nematode C. elegans

> *The future of molecular biology lies in the extension of research to other areas of biology, notably development and the nervous system*—Sydney Brenner, 1963

Brenner chose *C. elegans* because of its rapid life cycle, fecundity, genetic tractability, and simple cellular complexity of only ~1,000 cells. The first five years of his research focused on genetics of *C. elegans*, resulting in isolation and characterization of ~100 genes by methanesulfonic acid, ethyl ester (EMS) mutagenesis and screening for visible phenotypes (10). In the next decade, John Sulston, Bob Horvitz, and colleagues proceeded to observe and describe the complete lineage, from fertilized egg to adult, of both male and hermaphrodite, detailed in a heroic series of papers (81–84). Determination of "The Mind of the Worm" with the entire reconstruction of the hermaphrodite nervous system, an equally ambitious undertaking, was published in 1986 (99). Reconstruction of the male nervous system is in progress (Scott Emmons and David Hall, personal communication). With knowledge of the entire cell lineage in hand, cellular function can be determined by selective cell ablation with a laser microbeam (85). The cellular basis of numerous developmental and

behavioral processes has been ascertained using this approach. The powerful classic, reverse, and molecular genetic tools in *C. elegans* have enabled the study of basic biological problems at the cellular, genetic, molecular, and biochemical levels (50, 80). *C. elegans* has proven extremely useful for defining pathways of gene action, for identification of new proteins involved in a particular pathway, and for modeling the molecular mechanisms of human disease. For his work on the molecular genetic basis of programmed cell death, Bob Horvitz shared the 2002 Nobel Prize in Medicine or Physiology with Brenner and Sulston for their pioneering accomplishments in *C. elegans*.

The *C. elegans* genome was the first multicellular organism to be completely sequenced, an effort spearheaded by John Sulston, Robert Waterston, and Alan Coulson (12). About 50% of *C. elegans* genes are novel and do not share similarity to genes of organisms outside the Nematoda phylum. As many nematodes are plant or animal parasites, these nematode-specific genes represent excellent drug targets for prevention and treatment of nematode pathogens. Another 43% of *C. elegans* genes have human homologs, including numerous disease genes (22). *C. elegans* may be an effective organism for studying basic molecular genetic mechanisms underlying human disorders. Autosomal dominant polycystic kidney disease (ADPKD) affects 1 in 1,000 individuals, with mutation in either of two loci, PKD1 or PKD2, accounting for greater than 95% of all cases (78). *lov-1* and *pkd-2* are the *C. elegans* homologs of PKD1 and PKD2 (4, 5). LOV-1 and PKD-2 localize to male-specific sensory cilia and are required for the male mating behaviors. Stunningly, the PKD gene products may serve an evolutionarily conserved function in cilia. Most recently, mammalian polycystins 1 and 2 (encoded by PKD1 and PKD2) have been demonstrated to localize to primary kidney cilia (65, 104) and PKD2 was implicated in function of nodal cilia (required for left/right axis determination) (65, 66). Integrating information obtained from animal models as diverse as worms and mice into testable hypotheses regarding polycystin function provides an exquisite example of the power of model organisms to unravel complex biological pathways.

Forward and reverse genetics in C. elegans. Forward genetics identifies genes required for a particular biological function (50). Animals are mutagenized and progeny are screened for a visible mutant phenotype. Conversely, a reverse genetics approach moves from known gene sequence to gene function using a battery of techniques to determine cellular, molecular, and physiological roles (Fig. 1). Each has distinct advantages and disadvantages. With forward genetics, positional cloning of the mutated gene is often time-consuming and the rate-limiting step. On the other hand, when starting with a phenotype of interest, gene function and pathways may be ascertained without bias. In a reverse genetics approach, the molecular identity of the gene is already known. However, knocking out or knocking down gene function does not guarantee a mutant phenotype. Redundant genes or subtle mutant defects may be overlooked.

To understand the biological function of a gene, one needs to know the site and timing of gene action (spatial and temporal expression) and the phenotypic consequences of altering or removing the gene. Two techniques developed to study gene expression and function in *C. elegans* have revolutionized biology: RNA interference (RNAi) to knock down gene function (29) and GFP as an expression marker (15).

RNAi. In *C. elegans*, lack of methodologies for knocking out gene function via homologous recombination, a common tool employed by a yeast geneticist, presented a huge obstacle. In response to this shortcoming, Fire and Mello (29, 88, 89) developed double-stranded RNA (dsRNA)-dependent posttranscriptional gene silencing, or RNAi. This technology revolutionized the *C. elegans* field, as illustrated by genome-wide RNAi screens (27, 32, 57), and extends its power to a variety of organisms, ranging from *Trypanosoma brucei* (76) to mammals, most recently human cells (56, 59, 64). Excellent recent reviews about RNAi are available and will not be repeated here (40, 41, 87).

dsRNA may be experimentally introduced to *C. elegans* in several ways: by direct injection into the hermaphrodite germ line, by soaking in dsRNA (58), by ingestion of bacteria engineered to produce dsRNA (88, 89), or by generating heritable inverted repeat (IR) genes (86). The advantages of the transgenic IR RNAi method include the inactivation of genes that act in the nervous system (a tissue that is typically resistant to RNAi), the generation of a large number of RNAi mutants, and the production of easily maintained transgenic RNAi strains. A major drawback of transgenic IR RNAi is that it is not a technique easily amenable to high-throughput screening.

Genome-wide RNAi screens have been executed using injected, soaked, or ingested dsRNA. Injected RNAi of genes on chromosome III coupled with differential interference contrast and dissecting microscopy revealed a cell division function for 13% of over 2,200 ORFs in total (38). RNAi feeding libraries that inactivate most of the annotated *C. elegans* genes on chromosomes I and II (32, 57) are available for a small fee and have been used to identify genes required for fertility, embryonic and postembryonic development, movement, and longevity (26, 32, 57). It is only a matter of time until RNAi libraries exist for all six chromosomes and RNAi phenotypes are cataloged for all annotated *C. elegans* genes. It is also very likely that the first completed functional genome of a multicellular organism will be that of The Worm.

Gene expression analysis in C. elegans. Two main approaches are employed to examine *C. elegans* gene expression patterns: reporter gene constructs and microarrays. Use of GFP as a fluorescent reporter for gene expression or subcellular protein localization has become standard, proceeds on a gene-by-gene basis, and relies largely on the cell identification ability of the microscopist. Hence, this approach is not easily upscaled, and efforts are being made to compile and make

available expression data. In contrast, microarray technology now enables *C. elegans* researchers to investigate and model complexity. Creative uses of *C. elegans* microarrays have provided new approaches for tissue-specific gene profiling, for examining genome organization, and for exploring global changes in gene expression (for review, see Ref. 72).

A WORM'S TOUCH. An excellent example of comprehensive of techniques available in *C. elegans* to dissect the molecular basis of behavior comes from the work of Marty Chalfie and colleagues (13, 14, 39). Chalfie initiated his studies on touch sensitivity, or mechanosensation, by laser ablation of microtubule cells and assaying behavioral responses of operated animals. Animals lacking touch cells fail to respond to gentle body touch (13, 14). Taking a forward genetics approach and looking for mutants defective in mechanosensation (or Mec), they identified genes required for touch receptor development, differentiation, and function (for review, see Ref. 27). Cloning of the genes required for touch cell function revealed an evolutionarily conserved mechanosensitive channel (for review, see Ref. 35). MEC channel activity was demonstrated in *Xenopus* oocytes (39). Most recently, cDNA microarrays were used to study gene expression profiles for *C. elegans* touch receptor neurons. This functional genomics approach successfully identified new *mec* genes that were not identified in conventional genetic screens as well as providing candidates for uncloned previously identified *mec* genes (107).

Summary

This article has discussed the use of model organisms and the current technologies to dissect gene function at both individual and genome-wide levels. Gene annotation and assigning meaning and function to genome sequence remains a challenge. Every genetic model system has a greater number of genes than mutants, making high-throughput methods for generating knockouts a necessity. Adapting the genome-wide "bar-coded" approach in *S. cerevisiae* or RNAi screens in *C. elegans* to other organisms is highly promising.

There is more to life than a DNA sequence. Development of genomic and genetic resources for other model systems is a worthy endeavor. Sound investments in the model organism portfolio might include well-established developmental systems such as the ascidian *Ciona*, the frogs *X. laevis* and *X. tropicalis*, *Gallus gallus* (chicken), and *Rattus norvegicus* (rat), cell biological models including *Strongylocentrotus purpuratus* (sea urchin) and *Chlamydomonas reinhartii* (a unicellular alga), animal parasites such as *P. falciparum* (malaria), and the crop plant *Zea mays* (maize).

How to interpret sequence information to understanding of gene and protein function and pathways, cellular function, physiology, and the generation of an organism will keep modern scientists busy for quite a while. Simply identifying genes in the 3,000-Mb human genome is limited by current gene prediction methods and remains a major challenge. Many human genes have homologs in model organisms as simple as bacteria or single-celled eukaryotes, suggesting a conserved function. However, the role of most predicted genes remains unknown. Gene functions will be revealed through the powerful molecular genetic tools available in model organism coupled with emerging high-throughput technologies. Gene identification will be aided by comparative genomics with a smaller and less complex genome. Model organisms will expedite gene annotation and serve as a beacon for understanding how genes specify an organism and how gene perturbations may lead to human disease.

I am especially grateful to Dr. David R. Sherwood and the two anonymous reviewers for comments on this manuscript. I also thank the members of my laboratory for ongoing useful discussions. Work in my laboratory is supported by grants from the National Institutes of Health, the PKD Foundation, and the American Heart Association.

REFERENCES

1. **Adams MD et al. (Celera Genomics).** The genome sequence of *Drosophila melanogaster*. *Science* 287: 2185–2195, 2000.
2. **Aparicio S, Chapman J, Stupka E, Putnam N, Chia JM, Dehal P, Christoffels A, Rash S, Hoon S, Smit A, Gelpke MD, Roach J, Oh T, Ho IY, Wong M, Detter C, Verhoef F, Predki P, Tay A, Lucas S, Richardson P, Smith SF, Clark MS, Edwards YJ, Doggett N, Zharkikh A, Tavtigian SV, Pruss D, Barnstead M, Evans C, Baden H, Powell J, Glusman G, Rowen L, Hood L, Tan YH, Elgar G, Hawkins T, Venkatesh B, Rokhsar D, and Brenner S.** Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* 297: 1301–1310, 2002.
3. **Arabidopsis Genome Initiative.** Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796–815, 2000.
4. **Barr MM, DeModena J, Braun D, Nguyen CQ, Hall DH, and Sternberg PW.** The *Caenorhabditis elegans* autosomal dominant polycystic kidney disease gene homologs lov-1 and pkd-2 act in the same pathway. *Curr Biol* 11: 1341–1346, 2001.
5. **Barr MM and Sternberg PW.** A polycystic kidney-disease gene homologue required for male mating behaviour in *C. elegans*. *Nature* 401: 386–389, 1999.
6. **Beggs JD.** Transformation of yeast by a replicating hybrid plasmid. *Nature* 275: 104–109, 1978.
7. **Birrell GW, Giaever G, Chu AM, Davis RW, and Brown JM.** A genome-wide screen in *Saccharomyces cerevisiae* for genes affecting UV radiation sensitivity. *Proc Natl Acad Sci USA* 98: 12608–12613, 2001.
8. **Blaxter M.** *Caenorhabditis elegans* is a nematode. *Science* 282: 2041–2046, 1998.
9. **Brem RB, Yvert G, Clinton R, and Kruglyak L.** Genetic dissection of transcriptional regulation in budding yeast. *Science* 296: 752–755, 2002.
10. **Brenner S.** The genetics of *Caenorhabditis elegans*. *Genetics* 77: 71–94, 1974.
11. **Brenner S, Elgar G, Sandford R, Macrae A, Venkatesh B, and Aparicio S.** Characterization of the pufferfish (Fugu) genome as a compact model vertebrate genome. *Nature* 366: 265–268, 1993.
12. *C. elegans* Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 282: 2012–2018, 1998.
13. **Chalfie M and Sulston J.** Developmental genetics of the mechanosensory neurons of *Caenorhabditis elegans*. *Dev Biol* 82: 358–370, 1981.
14. **Chalfie M, Sulston JE, White JG, Southgate E, Thomson JN, and Brenner S.** The neural circuit for touch sensitivity in *Caenorhabditis elegans*. *J Neurosci* 5: 956–964, 1985.

15. **Chalfie M, Tu Y, Euskirchen G, Ward WW, and Prasher DC.** Green fluorescent protein as a marker for gene expression. *Science* 263: 802–805, 1994.
16. **Cho RJ, Campbell MJ, Winzeler EA, Steinmetz L, Conway A, Wodicka L, Wolfsberg TG, Gabrielian AE, Landsman D, Lockhart DJ, and Davis RW.** A genome-wide transcriptional analysis of the mitotic cell cycle. *Mol Cell* 2: 65–73, 1998.
17. **Christensen M, Estevez A, Yin X, Fox R, Morrison R, McDonnell M, Gleason C, Miller DM III, and Strange K.** A primary culture system for functional analysis of *C. elegans* neurons and muscle cells. *Neuron* 33: 503–514, 2002.
18. **Christophides GK, Zdobnov E, Barillas-Mury C, Birney E, Blandin S, Blass C, Brey PT, Collins FH, Danielli A, Dimopoulos G, Hetru C, Hoa NT, Hoffmann JA, Kanzok SM, Letunic I, Levashina EA, Loukeris TG, Lycett G, Meister S, Michel K, Moita LF, Muller HM, Osta MA, Paskewitz SM, Reichhart JM, Rzhetsky A, Troxler L, Vernick KD, Vlachou D, Volz J, von Mering C, Xu J, Zheng L, Bork P, and Kafatos FC.** Immunity-related genes and gene families in *Anopheles gambiae*. *Science* 298: 159–165, 2002.
19. **Chu S, DeRisi J, Eisen M, Mulholland J, Botstein D, Brown PO, and Herskowitz I.** The transcriptional program of sporulation in budding yeast. *Science* 282: 699–705, 1998.
20. **Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE III, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Barrell BG, et al.** Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393: 537–544, 1998.
21. **Costanzo MC, Crawford ME, Hirschman JE, Kranz JE, Olsen P, Robertson LS, Skrzypek MS, Braun BR, Hopkins KL, Kondu P, Lengieza C, Lew-Smith JE, Tillberg M, and Garrels JI.** YPD, PombePD and WormPD: model organism volumes of the BioKnowledge library, an integrated resource for protein information. *Nucleic Acids Res* 29: 75–79, 2001.
22. **Culetto E and Sattelle DB.** A role for *Caenorhabditis elegans* in understanding the function and interactions of human disease genes. *Hum Mol Genet* 9: 869–877, 2000.
23. **Daly MJ.** Estimating the human gene count. *Cell* 109: 283–284, 2002.
24. **Das M, Burge CB, Park E, Colinas J, and Pelletier J.** Assessment of the total number of human transcription units. *Genomics* 77: 71–78, 2001.
25. **DeRisi JL, Iyer VR, and Brown PO.** Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278: 680–686, 1997.
26. **Dillin A, Hsu AL, Arantes-Oliveira N, Lehrer-Graiwer J, Hsin H, Fraser AG, Kamath RS, Ahringer J, and Kenyon C.** Rates of behavior and aging specified by mitochondrial function during development. *Science* 298: 2398–2401, 2002.
27. **Driscoll M and Tavernarakis N.** Molecules that mediate touch transduction in the nematode *Caenorhabditis elegans*. *Gravit Space Biol Bull* 10: 33–42, 1997.
28. **Fields S and Song O.** A novel genetic system to detect protein-protein interactions. *Nature* 340: 245–246, 1989.
29. **Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, and Mello CC.** Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391: 806–811, 1998.
30. **Fleischmann RD, Alland D, Eisen JA, Carpenter L, White O, Peterson J, DeBoy R, Dodson R, Gwinn M, Haft D, Hickey E, Kolonay JF, Nelson WC, Umayam LA, Ermolaeva M, Salzberg SL, Delcher A, Utterback T, Weidman J, Khouri H, Gill J, Mikula A, Bishai W, Jacobs WR Jr, Venter JC, and Fraser CM.** Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. *J Bacteriol* 184: 5479–5490, 2002.
31. **Forsburg SL.** The art and design of genetic screens: yeast. *Nat Rev Genet* 2: 659–668, 2001.
32. **Fraser AG, Kamath RS, Zipperlen P, Martinez-Campos M, Sohrmann M, and Ahringer J.** Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. *Nature* 408: 325–330, 2000.
33. **Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan MS, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Pertea M, Allen J, Selengut J, Haft D, Mather MW, Vaidya AB, Martin DM, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph SA, McFadden GI, Cummings LM, Subramanian GM, Mungall C, Venter JC, Carucci DJ, Hoffman SL, Newbold C, Davis RW, Fraser CM, and Barrell B.** Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419: 498–511, 2002.
34. **Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM, Remor M, Hofert C, Schelder M, Brajenovic M, Ruffner H, Merino A, Klein K, Hudak M, Dickson D, Rudi T, Gnau V, Bauch A, Bastuck S, Huhse B, Leutwein C, Heurtier MA, Copley RR, Edelmann A, Querfurth E, Rybin V, Drewes G, Raida M, Bouwmeester T, Bork P, Seraphin B, Kuster B, Neubauer G, and Superti-Furga G.** Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415: 141–147, 2002.
35. **Gillespie PG and Walker RG.** Molecular basis of mechanosensory transduction. *Nature* 413: 194–202, 2001.
36. **Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A, and Briggs S.** A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* 296: 92–100, 2002.
37. **Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, and Oliver SG.** Life with 6000 genes. *Science* 274: 546, 563–567, 1996.
38. **Gonczy P, Echeverri G, Oegema K, Coulson A, Jones SJ, Copley RR, Duperon J, Oegema J, Brehm M, Cassin E, Hannak E, Kirkham M, Pichler S, Flohrs K, Goessen A, Leidel S, Alleaume AM, Martin C, Ozlu N, Bork P, and Hyman AA.** Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. *Nature* 408: 331–336, 2000.
39. **Goodman MB, Ernstrom GG, Chelur DS, O'Hagan R, Yao CA, and Chalfie M.** MEC2 regulates *C. elegans* DEG/ENaC channels needed for mechanosensation. *Nature* 415: 1039–1042, 2002.
40. **Grishok A and Mello CC.** RNAi (Nematodes: *Caenorhabditis elegans*). *Adv Genet* 46: 339–360, 2002.
41. **Hannon GJ.** RNA interference. *Nature* 418: 244–251, 2002.
42. **Hartwell LH, Culotti J, and Reid B.** Genetic control of the cell-division cycle in yeast. I. Detection of mutants. *Proc Natl Acad Sci USA* 66: 352–359, 1970.
43. **Ho Y, Gruhler A, Heilbut A, Bader GD, Moore L, Adams SL, Millar A, Taylor P, Bennett K, Boutilier K, Yang L, Wolting C, Donaldson I, Schandorff S, Shewnarane J, Vo M, Taggart J, Goudreault M, Muskat B, Alfarano C, Dewar D, Lin Z, Michalickova K, Willems AR, Sassi H, Nielsen PA, Rasmussen KJ, Andersen JR, Johansen LE, Hansen LH, Jespersen H, Podtelejnikov A, Nielsen E, Crawford J, Poulsen V, Sorensen BD, Matthiesen J, Hendrickson RC, Gleeson F, Pawson T, Moran MF, Durocher D, Mann M, Hogue CW, Figeys D, and Tyers M.** Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* 415: 180–183, 2002.

44. **Hogenesch JB, Ching KA, Batalov S, Su AI, Walker JR, Zhou Y, Kay SA, Schultz PG, and Cooke MP.** A comparison of the Celera and Ensembl predicted gene sets reveals little overlap in novel genes. *Cell* 106: 413–415, 2001.
45. **Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, Nusskern DR, Wincker P, Clark AG, Ribeiro JM, Wides R, Salzberg SL, Loftus B, Yandell M, Majoros WH, Rusch DB, Lai Z, Kraft CL, Abril JF, Anthouard V, Arensburger P, Atkinson PW, Baden H, de Berardinis V, Baldwin D, Benes V, Biedler J, Blass C, Bolanos R, Boscus D, Barnstead M, Cai S, Center A, Chatuverdi K, Christophides GK, Chrystal MA, Clamp M, Cravchik A, Curwen V, Dana A, Delcher A, Dew I, Evans CA, Flanigan M, Grundschober-Freimoser A, Friedli L, Gu Z, Guan P, Guigo R, Hillenmeyer ME, Hladun SL, Hogan JR, Hong YS, Hoover J, Jaillon O, Ke Z, Kodira C, Kokoza E, Koutsos A, Letunic I, Levitsky A, Liang Y, Lin JJ, Lobo NF, Lopez JR, Malek JA, McIntosh TC, Meister S, Miller J, Mobarry C, Mongin E, Murphy SD, O'Brochta DA, Pfannkoch C, Qi R, Regier MA, Remington K, Shao H, Sharakhova MV, Sitter CD, Shetty J, Smith TJ, Strong R, Sun J, Thomasova D, Ton LQ, Topalis P, Tu Z, Unger MF, Walenz B, Wang A, Wang J, Wang M, Wang X, Woodford KJ, Wortman JR, Wu M, Yao A, Zdobnov EM, Zhang H, Zhao Q, et al.** The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* 298: 129–149, 2002.
47. **Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M, and Sakaki Y.** A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci USA* 98: 4569–4574, 2001.
48. **Jelinsky SA, Estep P, Church GM, and Samson LD.** Regulatory networks revealed by transcriptional profiling of damaged *Saccharomyces cerevisiae* cells: Rpn4 links base excision repair with proteasomes. *Mol Cell Biol* 20: 8157–8167, 2000.
49. **Jelinsky SA and Samson LD.** Global response of *Saccharomyces cerevisiae* to an alkylating agent. *Proc Natl Acad Sci USA* 96: 1486–1491, 1999.
50. **Jorgensen EM and Mango SE.** The art and design of genetic screens: *Caenorhabditis elegans*. *Nat Rev Genet* 3: 356–369, 2002.
51. **Kumar A, Agarwal S, Heyman JA, Matson S, Heidtman M, Piccirillo S, Umansky L, Drawid A, Jansen R, Liu Y, Cheung KH, Miller P, Gerstein M, Roeder GS, and Snyder M.** Subcellular localization of the yeast proteome. *Genes Dev* 16: 707–719, 2002.
52. **Kumar A, Harrison PM, Cheung KH, Lan N, Echols N, Bertone P, Miller P, Gerstein MB, and Snyder M.** An integrated approach for finding overlooked genes in yeast. *Nat Biotechnol* 20: 58–63, 2002.
53. **Kumar A and Snyder M.** Emerging technologies in yeast genomics. *Nat Rev Genet* 2: 302–312, 2001.
54. **Kumar A and Snyder M.** Protein complexes take the bait. *Nature* 415: 123–124, 2002.
55. **Lander ES et al. (International Human Genome Sequencing Consortium).** Initial sequencing and analysis of the human genome. *Nature* 409: 860–921, 2001.
56. **Lee NS, Dohjima T, Bauer G, Li H, Li MJ, Ehsani A, Salvaterra P, and Rossi J.** Expression of small interfering RNAs targeted against HIV-1 rev transcripts in human cells. *Nat Biotechnol* 20: 500–505, 2002.
57. **Lee SS, Lee RY, Fraser AG, Kamath RS, Ahringer J, and Ruvkun G.** A systematic RNAi screen identifies a critical role for mitochondria in *C. elegans* longevity. *Nat Genet* 25: 25, 2002.
58. **Maeda I, Kohara Y, Yamamoto M, and Sugimoto A.** Large-scale analysis of gene function in *Caenorhabditis elegans* by high-throughput RNAi. *Curr Biol* 11: 171–176, 2001.
59. **Miyagishi M and Taira K.** U6 promoter driven siRNAs with four uridine 3′ overhangs efficiently suppress targeted gene expression in mammalian cells. *Nat Biotechnol* 20: 497–500, 2002.
60. **Murray AW and Marks D.** Can sequencing shed light on cell cycling? *Nature* 409: 844–846, 2001.
60a. **National Institutes of Health.** Trans-NIH *Xenopus* Initiative. *NIH Model Organisms for Biomedical Research* [Online]. http://www.nih.gov/science/models/xenopus/ [2000].
61. **Nurse P.** Genetic control of cell size at cell division in yeast. *Nature* 256: 547–551, 1975.
62. **Nurse P.** Universal control mechanism regulating onset of M-phase. *Nature* 344: 503–508, 1990.
63. **Oliver S.** "To-day, we have naming of parts. . ." *Nat Biotechnol* 20: 27–28, 2002.
64. **Paul CP, Good PD, Winer I, and Engelke DR.** Effective expression of small interfering RNA in human cells. *Nat Biotechnol* 20: 505–508, 2002.
65. **Pazour GJ, San Agustin JT, Follit JA, Rosenbaum JL, and Witman GB.** Polycystin-2 localizes to kidney cilia and the ciliary level is elevated in orpk mice with polycystic kidney disease. *Curr Biol* 12: R378–R380, 2002.
66. **Pennekamp P, Karcher C, Fischer A, Schweickert A, Skyrabin B, Horst J, Blum M, and Dworniczak B.** The ion channel polycystin-2 is required for left-right axis determination in mice. *Curr Biol* 12: 938–943, 2002.
67. **Petropoulos CJ.** *Retroviruses*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1997.
68. **Primig M, Williams RM, Winzeler EA, Tevzadze GG, Conway AR, Hwang SY, Davis RW, and Esposito RE.** The core meiotic transcriptome in budding yeasts. *Nat Genet* 26: 415–423, 2000.
70. **Raghuraman MK, Winzeler EA, Collingwood D, Hunt S, Wodicka L, Conway A, Lockhart DJ, Davis RW, Brewer BJ, and Fangman WL.** Replication dynamics of the yeast genome. *Science* 294: 115–121, 2001.
71. **Ranson H, Claudianos C, Ortelli F, Abgrall C, Hemingway J, Sharakhova MV, Unger MF, Collins FH, and Feyereisen R.** Evolution of supergene families associated with insecticide resistance. *Science* 298: 179–181, 2002.
72. **Reinke V.** Functional exploration of the *C. elegans* genome using DNA microarrays. *Nat Genet* 32, *Suppl* 2: 541–546, 2002.
73. **Riehle MA, Garczynski SF, Crim JW, Hill CA, and Brown MR.** Neuropeptides and peptide hormones in *Anopheles gambiae*. *Science* 298: 172–175, 2002.
74. **Rothstein RJ.** One-step gene disruption in yeast. *Methods Enzymol* 101: 202–211, 1983.
75. **Schneider D and Shahabuddin M.** Malaria parasite development in a *Drosophila* model. *Science* 288: 2376–2379, 2000.
76. **Shi H, Djikeng A, Mark T, Wirtz E, Tschudi C, and Ullu E.** Genetic interference in *Trypanosoma brucei* by heritable and inducible double-stranded RNA. *RNA* 6: 1069–1076, 2000.
77. **Shoemaker DD, Lashkari DA, Morris D, Mittmann M, and Davis RW.** Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar-coding strategy. *Nat Genet* 14: 450–456, 1996.
78. **Somlo S and Ehrlich B.** Human disease: calcium signaling in polycystic kidney disease. *Curr Biol* 11: R356–R360, 2001.
79. **Spellman PT, Sherlock G, Zhang MQ, Iyer VR, Anders K, Eisen MB, Brown PO, Botstein D, and Futcher B.** Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol Biol Cell* 9: 3273–3297, 1998.
80. **Sternberg PW.** Working in the post-genomic *C. elegans* world. *Cell* 105: 173–176, 2001.
81. **Sulston JE.** Post-embryonic development in the ventral cord of *Caenorhabditis elegans*. *Philos Trans R Soc Lond B Biol Sci* 275: 287–297, 1976.
82. **Sulston JE, Albertson DG, and Thomson JN.** The *Caenorhabditis elegans* male: postembryonic development of nongonadal structures. *Dev Biol* 78: 542–576, 1980.
83. **Sulston JE and Horvitz HR.** Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev Biol* 56: 110–156, 1977.
84. **Sulston JE, Schierenberg E, White JG, and Thomson JN.** The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev Biol* 100: 64–119, 1983.
85. **Sulston JE and White JG.** Regulation and cell autonomy during postembryonic development of *Caenorhabditis elegans*. *Dev Biol* 78: 577–597, 1980.

86. **Tavernarakis N, Wang SL, Dorovkov M, Ryazanov A, and Driscoll M.** Heritable and inducible genetic interference by double-stranded RNA encoded by transgenes. *Nat Genet* 24: 180–183, 2000.
87. **Tijsterman M, Ketting RF, and Plasterk RH.** The genetics of RNA silencing. *Annu Rev Genet* 36: 489–519, 2002.
88. **Timmons L, Court DL, and Fire A.** Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. *Gene* 263: 103–112, 2001.
89. **Timmons L and Fire A.** Specific interference by ingested dsRNA. *Nature* 395: 854, 1998.
90. **Tong AH, Drees B, Nardelli G, Bader GD, Brannetti B, Castagnoli L, Evangelista M, Ferracuti S, Nelson B, Paoluzi S, Quondam M, Zucconi A, Hogue CW, Fields S, Boone C, and Cesareni G.** A combined experimental and computational strategy to define protein interaction networks for peptide recognition modules. *Science* 295: 321–324, 2002.
91. **Tong AH, Evangelista M, Parsons AB, Xu H, Bader GD, Page N, Robinson M, Raghibizadeh S, Hogue CW, Bussey H, Andrews B, Tyers M, and Boone C.** Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* 294: 2364–2368, 2001.
92. **Uetz P.** Two-hybrid arrays. *Curr Opin Chem Biol* 6: 57–62, 2002.
93. **Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P, Qureshi-Emili A, Li Y, Godwin B, Conover D, Kalbfleisch T, Vijayadamodar G, Yang M, Johnston M, Fields S, and Rothberg JM.** A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* 403: 623–627, 2000.
94. **Venkatesh B, Gilligan P, and Brenner S.** Fugu: a compact vertebrate reference genome. *FEBS Lett* 476: 3–7, 2000.
95. **Venter JC et al. (Celera Genomics).** The sequence of the human genome. *Science* 291: 1304–1351, 2001.
96. **Von Mering C, Krause R, Snel B, Cornell M, Oliver SG, Fields S, and Bork P.** Comparative assessment of large-scale data sets of protein-protein interactions. *Nature* 417: 399–403, 2002.
97. **Waterston RH et al. (Mouse Genome Sequencing Consortium).** Initial sequencing and comparative analysis of the mouse genome. *Nature* 420: 520–562, 2002.
98. **Weinstein DC and Hemmati-Brivanlou A.** Neural induction. *Annu Rev Cell Dev Biol* 15: 411–433, 1999.
99. **White JG, Southgate E, Thomson JN, and Brenner S.** The structure of the nervous system of the nematode *Caenorhabditis elegans*: the mind of a worm. *Philos Trans R Soc Lond* 314: 1–340, 1986.
100. **Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Chu AM, Connelly C, Davis K, Dietrich F, Dow SW, El Bakkoury M, Foury F, Friend SH, Gentalen E, Giaever G, Hegemann JH, Jones T, Laub M, Liao H, Davis RW, et al.** Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* 285: 901–906, 1999.
101. **Wood V, Gwilliam R, Rajandream MA, Lyne M, Lyne R, Stewart A, Sgouros J, Peat N, Hayles J, Baker S, Basham D, Bowman S, Brooks K, Brown D, Brown S, Chillingworth T, Churcher C, Collins M, Connor R, Cronin A, Davis P, Feltwell T, Fraser A, Gentles S, Goble A, Hamlin N, Harris D, Hidalgo J, Hodgson G, Holroyd S, Hornsby T, Howarth S, Huckle EJ, Hunt S, Jagels K, James K, Jones L, Jones M, Leather S, McDonald S, McLean J, Mooney P, Moule S, Mungall K, Murphy L, Niblett D, Odell C, Oliver K, O'Neil S, Pearson D, Quail MA, Rabbinowitsch E, Rutherford K, Rutter S, Saunders D, Seeger K, Sharp S, Skelton J, Simmonds M, Squares R, Squares S, Stevens K, Taylor K, Taylor RG, Tivey A, Walsh S, Warren T, Whitehead S, Woodward J, Volckaert G, Aert R, Robben J, Grymonprez B, Weltjens I, Vanstreels E, Rieger M, Schafer M, Muller-Auer S, Gabel C, Fuchs M, Fritzc C, Holzer E, Moestl D, Hilbert H, Borzym K, Langer I, Beck A, Lehrach H, Reinhardt R, Pohl TM, Eger P, Zimmermann W, Wedler H, Wambutt R, Purnelle B, Goffeau A, Cadieu E, Dreano S, Gloux S, Lelaure V, Mottier S, Galibert F, Aves SJ, Xiang Z, Hunt C, Moore K, Hurst SM, Lucas M, Rochet M, Gaillardin C, Tallada VA, Garzon A, Thode G, Daga RR, Cruzado L, Jimenez J, Sanchez M, del Rey F, Benito J, Dominguez A, Revuelta JL, Moreno S, Armstrong J, Forsburg SL, Cerutti L, Lowe T, McCombie WR, Paulsen I, Potashkin J, Shpakovski GV, Ussery D, Barrell BG, Nurse P, and Cerrutti L.** The genome sequence of *Schizosaccharomyces pombe*. *Nature* 415: 871–880, 2002.
102. **Wright FA, Lemon WJ, Zhao WD, Sears R, Zhuo D, Wang JP, Yang HY, Baer T, Stredney D, Spitzner J, Stutz A, Krahe R, and Yuan B.** A draft annotation and overview of the human genome. *Genome Biol* 2: 2001.
103. **Wyrick JJ, Aparicio JG, Chen T, Barnett JD, Jennings EG, Young RA, Bell SP, and Aparicio OM.** Genome-wide distribution of ORC and MCM proteins in *S. cerevisiae*: high-resolution mapping of replication origins. *Science* 294: 2357–2360, 2001.
104. **Yoder BK, Hou X, and Guay-Woodford LM.** The polycystic kidney disease proteins, polycystin-1, polycystin-2, polaris, and cystin, are co-localized in renal cilia. *J Am Soc Nephrol* 13: 2508–2516, 2002.
105. **Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M, Liu J, Sun J, Tang J, Chen Y, Huang X, Lin W, Ye C, Tong W, Cong L, Geng J, Han Y, Li L, Li W, Hu G, Li J, Liu Z, Qi Q, Li T, Wang X, Lu H, Wu T, Zhu M, Ni P, Han H, Dong W, Ren X, Feng X, Cui P, Li X, Wang H, Xu X, Zhai W, Xu Z, Zhang J, He S, Xu J, Zhang K, Zheng X, Dong J, Zeng W, Tao L, Ye J, Tan J, Chen X, He J, Liu D, Tian W, Tian C, Xia H, Bao Q, Li G, Gao H, Cao T, Zhao W, Li P, Chen W, Zhang Y, Hu J, Liu S, Yang J, Zhang G, Xiong Y, Li Z, Mao L, Zhou C, Zhu Z, Chen R, Hao B, Zheng W, Chen S, Guo W, Tao M, Zhu L, Yuan L, and Yang H.** A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* 296: 79–92, 2002.
106. **Zeng Q, Morales AJ, and Cottarel G.** Fungi and humans: closer than you think. *Trends Genet* 17: 682–684, 2001.
107. **Zhang Y, Ma C, Delohery T, Nasipak B, Foat BC, Bounoutas A, Bussemaker HJ, Kim SK, and Chalfie M.** Identification of genes expressed in *C. elegans* touch receptor neurons. *Nature* 418: 331–335, 2002.
108. **Zhu H, Bilgin M, Bangham R, Hall D, Casamayor A, Bertone P, Lan N, Jansen R, Bidlingmaier S, Houfek T, Mitchell T, Miller P, Dean RA, Gerstein M, and Snyder M.** Global analysis of protein activities using proteome chips. *Science* 293: 2101–2105, 2001.