Finding function in novel targets: *C. elegans* as a model organism

Titus Kaletta and Michael O. Hengartner‡*

Abstract | Despite its apparent simplicity, the nematode worm *Caenorhabditis elegans* has developed into an important model for biomedical research, particularly in the functional characterization of novel drug targets that have been identified using genomics technologies. The cellular complexity and the conservation of disease pathways between *C. elegans* and higher organisms, together with the simplicity and cost-effectiveness of cultivation, make for an effective ***in vivo*** model that is amenable to whole-organism high-throughput compound screens and large-scale target validation. This review describes how *C. elegans* models can be used to advance our understanding of the molecular mechanisms of drug action and disease pathogenesis.

In the second half of the twentieth century, Sydney Brenner introduced the soil nematode *Caenorhabditis elegans* as a model to study development and neurobiology[1]. Today, *C. elegans* is used to study a much larger variety of biological processes including apoptosis, cell signalling, cell cycle, cell polarity, gene regulation, metabolism, ageing and sex determination[2]. Many key discoveries, both in basic biology and medically relevant areas, were first made in the worm (BOX 1).

Together, these studies revealed a surprisingly strong conservation in molecular and cellular pathways between worms and mammals. Indeed, subsequent comparison of the human and *C. elegans* genomes confirmed that the majority of human disease genes and disease pathways are present in *C. elegans*[3].

C. elegans has a number of features that make it a powerful tool for the pharmaceutical industry. First, it is easy to culture: although the animal normally grows in the soil and feeds on various bacteria, it can readily be raised in the laboratory on a diet of *Escherichia coli*. Second, it reproduces rapidly and prolifically: within 3 days it develops from egg to an adult worm of 1.3 mm in length. Short generation time and about 300 progenies per self-fertilizing hermaphrodite enable the large-scale production of several million animals per day. Third, because of its small size, most assays can be carried out in microtitre plates either on agar or in liquid using more than one hundred animals in a single well of a 96-well plate. Fourth, the worm is transparent and, with the use of *in vivo* fluorescence markers, processes such as axon growth, embryogenesis and fat metabolism can easily be studied in the living animal. Fifth, it is a sophisticated multicellular animal: although the adult hermaphrodite has only 959 somatic cells, these form many different organs and tissues including muscle, hypodermis (skin), intestine, reproductive system, glands, and a nervous system containing 302 neurons[2,4].

Relevance of the *C. elegans* system

Although *C. elegans* has provided many insights into the underlying mechanisms of human diseases, many still question whether *C. elegans* can really be used as disease model and, if so, how relevant such a model can be. Does *C. elegans* demonstrate behaviours that could be characterized as mood disorders? Can worms be depressed? In most cases, there will not be a direct correspondence between human pathology and *C. elegans* phenotypes. Given that even mammalian models are often not reliably predictive of drug action in humans, it is — from a preclinical model perspective — unrealistic to expect an invertebrate system to give enough confidence to predict drug action and safety in humans. Non-mammalian model organisms will be typically used in early research and should deliver fast answers to a discovery problem, such as the function of a gene, or pioneer medical research to define novel therapeutic entry points. Of the animal models, *C. elegans* is certainly the fastest and most amenable to cost-effective medium/high-throughput technologies. *C. elegans* is a valuable disease model if the disease can be defined on a molecular basis. For example, if the underlying cause of depression is a defect in serotonergic signalling, a *C. elegans* model can be developed to study serotonergic signalling in detail. Such models

**Devgen, Technologie-park 30, B-9052 Gent, Belgium.*
‡Institute of Molecular Biology, University of Zurich, Winterhurerstrasse 190, 8057 Zurich, Switzerland.
Correspondence to T.K.
e-mail: titus.kaletta@devgen.com
doi:10.1038/nrd2031
Published online 21 April 2006

Box 1 | **Key biomedical discoveries enabled by *C. elegans* research**

Alzheimer's disease (AD). In 1993, the first presenilin was discovered in *C. elegans*[64]. Two years later, mutations in the human presenilin-1 gene were associated with early-onset familial AD[65,66]. A remarkable functional conservation between *C. elegans* and human could be demonstrated: expression of the human presenilin-1 in *C. elegans* could rescue neuronal deficiencies of *C. elegans sel-12* presenilin mutants[67,68]. *C. elegans* research has further advanced the understanding of AD by identifying the presenilins as components of the γ-secretase complex, an important target in AD[69].

Diabetes type 2. In 1997, genetic studies in *C. elegans* identified negative regulators of the insulin signalling pathway. One of these genes, *daf-16*, encodes the *C. elegans* orthologue of the forkhead transcription factor FOXO[70]. Five years later, FOXO loss-of-function was found to rescue the diabetic phenotype of insulin-resistant mice[71].

Depression. *C. elegans* is not only an established genetic model but can also be used to investigate the underlying mechanisms of whole-animal pharmacology. For example, the antidepressant fluoxetine has been shown to increase serotonergic signalling in *C. elegans* by inhibiting the *C. elegans* orthologue of the serotonin re-uptake transporter SERT[47]. This has stimulated a number of investigations to identify additional mode of actions of fluoxetine and to further elucidate the molecular mechanism of depression.

could then be pharmacologically validated and could enhance research around the serotonergic signalling network. The advantage of *C. elegans* over *in vitro* or cellular models is that instead of studying the isolated drug–receptor interaction, a functional serotonergic synapse can be studied within the context of a whole organism, and instead of measuring changes in a fluorescence read-out, for example, the actual behavioural response of the animal can be monitored (FIG. 1).

Indeed, the apparently simple body plan of *C. elegans* should not belie the high complexity seen at the cellular and physiological levels. If we consider the interaction of a drug with its target in a *C. elegans* assay, the following hurdles have to be overcome (FIG. 2). First, a drug needs to enter the body (sometimes this can be achieved through the skin, but in many cases the drug needs to be ingested to enter the intestine). It has to be bioavailable to pass through the gut membranes and to reach the target tissue. At the target site, a drug would have to alter the activity of the primary drug target and to overcome not only cellular but also tissue regulative mechanisms to finally cause a measurable pharmacological effect in *C. elegans*. Together, these requirements mean that only molecules with *in vivo* activity will have a chance to trigger a response in *C. elegans*. Notably, whole-animal assays also increase the chance of identifying synergistic or off-target effects, as the drug can simultaneously interact with multiple targets. In conclusion, a *C. elegans* assay has many features of whole-animal pharmacology in mammals.

Conservation between *C. elegans* and humans

Modelling a human disease in *C. elegans* is most straightforward if the drug target is conserved between the two species. Indeed, this is often the case: depending on the particular bioinformatics approach used, *C. elegans* homologues have been identified for 60–80% of human genes[3,5–8].

There are continued efforts to define for each human gene its corresponding orthologue(s) in various model organisms[9]. For example, the database OrthoDisease (see Further information) has applied an orthologue detection algorithm on 2,466 human disease genes derived from Online Mendelian Inheritance in Man. These genes identified 1,354 orthologue clusters in *Mus musculus*, 724 in *Drosophila melanogaster*, 533 in *C. elegans* and 153 in *E. coli*[10].

These genome-to-genome comparison tools greatly facilitate the search for orthologues in a model organism (BOX 2). On the basis of the degree of sequence conservation they provide an estimate of the likelihood of conservation of function between two genes. But evolution of function does not always parallel evolution of sequence and there are therefore a number of cases in which the initial genome-to-genome analysis is not sufficient. For example, the principal function of the $GABA_A$ (γ-aminobutyric acid A) receptor, a GABA-gated chloride channel, is to mediate the GABAergic signal via chloride transportation into the postsynaptic cell[11], but it is also an important anti-epileptic drug target. This property is shared by all $GABA_A$ receptors irrespective of the degree of sequence conservation. The mentioned orthologue databases did not identify an orthologue cluster for the *C. elegans* $GABA_A$ gene *unc-49* (sequence identity/similarity between UNC-49 (also known as T21C12.1) and human *GABRA6* is 34%/51%), which could have suggested that *C. elegans* is not a relevant model for GABA research. However, on the contrary, dissection of GABA signalling in *C. elegans* has been highly successful and uncovered, for example, the family of vesicular GABA transporters (the human vesicular inhibitory amino acid transporter and the *C. elegans unc-47* define an orthologue cluster in the InParanoid database (see Further information))[12,13].

By contrast, for certain large gene families, such as the nuclear hormone receptors, it is difficult to unambiguously assign all *C. elegans* orthologues, despite their reasonably high sequence conservation. Therefore, the degree of sequence conservation should not be used as the sole criterion for determining the relevance of the *C. elegans* model. It is recommended that additional studies are carried out: for example, expressing the human gene in *C. elegans* for functional replacement of the endogenous putative orthologue; testing for functional conservation of other known genes in the pathway; or testing known drugs or bio-tools for conservation of the mode of action. These studies will then allow estimating the predictive value of the *C. elegans* model.

Genome-wide RNAi for target identification

The standard approach to investigating the function of a gene is to alter its activity. Forward and reverse genetics have been successfully applied in the past for this purpose. With the completion of numerous genome sequences it became possible, and desirable, to determine the function of all predicted genes in an organism. Several approaches have been taken towards this goal. Attempts are being made to knock out all important druggable genes in mice, estimated at between 3,000 and 10,000 genes[14,15]. This obviously tedious approach becomes even more daunting if a gene function must be studied in an appropriate disease context. Crossing all mouse knockouts into a specific genetic background might be unfeasible

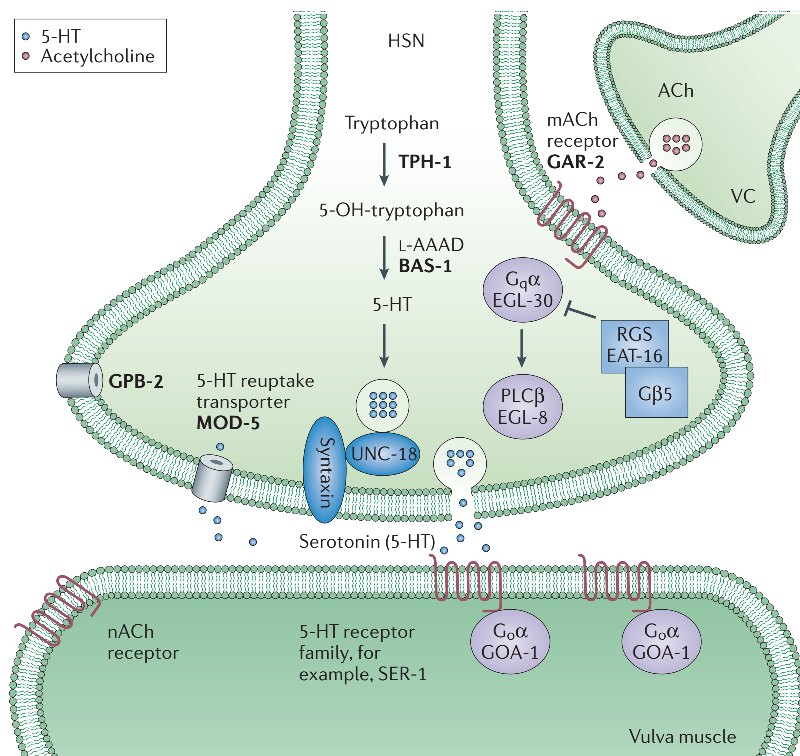

Figure 1 | **The serotonergic synapse of *C. elegans*.** A simplified cartoon of the hermaphrodite-specific neuron (HSN) neuromuscular junction. Two HSNs stimulate the 16 egg-laying muscles to contract and push embryos out through the vulva[81]. Egg-laying is stimulated by serotonergic signalling and modulated by cholinergic signalling[82–84]. The biochemical pathway for serotonin (5-hydroxytryptamine; 5-HT) synthesis is conserved between *C. elegans* and humans[85,86]. Docking of vesicles (grey spheres) requires the well-conserved SNARE complex (blue)[87]. Components of this complex (for example, UNC-18/nSec-1 and UNC-64/syntaxin), as well as many other proteins involved in vesicle exocytosis, were first identified through genetic screens in *C. elegans*[88,89]. Engagement of the serotonin receptor (for example, SER-1) on the postsynaptic cell activates a G-coupled cascade that stimulates muscle contraction. 5-HT is removed from the synapse by the 5-HT reuptake transporter MOD-5. HSN activity is modulated by the ventral-type C neurons[90]. These neurons release the neurotransmitter acetylcholine (ACh), which inhibits the muscarinic acetylcholine (mACh) receptor GAR-2 on the HSN neurons and consequently, via EGL-8 and EGL-30, inhibits egg-laying behaviour[91]. EAT-16 and GPB-2 integrate this signal to the $G_o\alpha - G_q\alpha$ signalling network[92]. BAS, biogenic amine synthesis related; EAT, eating (pharyngeal pumping) abnormal; EGL, egg-laying abnormal; GAR, G-protein-linked acetylcholine receptor; GOA, G-protein class O α-subunit; GPB, G-protein β-subunit; L-AAAD, L-aromatic amino acid decarboxylase; MOD, modulation of locomotion defective; VC, ventral-type C neuron.

or at least unaffordable. Another approach to studying the function of genes is to follow global patterns of gene expression using microarray technology. This high-throughput technology identifies genes that are up- or downregulated in a disease state but cannot reveal gene function or define causal roles for potential drug targets. Although a return to the 'old' positional cloning and linkage analysis will in individual cases be promising, a more systematic approach that uncovers all genes involved in a disease would be desirable, particularly if the disease has a multifactorial aetiology. RNA interference (RNAi), a phenomenon discovered in *C. elegans*[16], allows such systematic studies[17–19], and is therefore an ideal tool for high-throughput functional genomics[20].

RNAi-mediated knockdown of genes in *C. elegans* can be accomplished by directly injecting double-stranded RNA (dsRNA) into the animal, by soaking animals in a solution containing dsRNA, or by feeding *C. elegans* with bacteria that produce dsRNA[21,22]. Induction of RNAi in *C. elegans* requires dsRNA with a minimal fragment length of 100 base pairs. Typical genome-wide libraries use fragments in the range 500–1,500 base pairs[23,24]. In mammalian systems, only short 22-nucleotide dsRNA molecules (also known as small interfering RNA (siRNA)) are used, which helps to avoid an interferon response or nonspecific inhibition of protein synthesis through dsRNA-dependent protein kinases[25,26]. Neither response exists in *C. elegans*. The capacity to use long dsRNA in *C. elegans*, which gives rise to many different siRNA molecules, increases the efficiency of RNAi because the target mRNA gets attacked on many fronts. By contrast, this multitude of siRNAs might also increase the risk of inactivating other genes. Indeed, although RNAi has been assumed to cause a sequence specific knockdown, it has more recently become obvious that RNAi, at least *in vitro*, can also induce off-target effects[27]. However, systematic comparison of RNAi knockdowns with large mutant strain collections suggest that the threat posed by off-target effects might be minor in *C. elegans*, as most RNAi knockdown phenotypes can be reproduced with the corresponding gene knockout[28].

C. elegans animals exposed to dsRNA will develop a phenotype as soon as the activity of the gene product is significantly reduced. How long this takes will differ from gene to gene, and depends on the turnover rate of the encoded protein. Interestingly, RNAi can be induced at any time in the animal's life cycle, and therefore offers the opportunity to study gene function at all stages. At the other extreme, RNAi-induced phenotypes can be maintained over several generations simply by continuously feeding *C. elegans* on bacteria producing the relevant dsRNA.

The ease with which RNAi experiments can be performed, and the availability of the relevant genome data, have enabled a novel, high-throughput, systematic reverse genetics approach known as genome-wide RNAi (FIG. 3). In the late 1990s, several research institutes and companies such as Devgen and Exelixis built libraries of either *in vitro* synthesized dsRNA or bacteria that produce dsRNA[16–18,21,22,29]. These libraries, which represent a large fraction of all genes of the *C. elegans* genome, are miniaturized in 96-well plates and are therefore compatible with standard high-throughput robotics[20]. The 'whole genome' collection of *C. elegans* nicely fits in 210 96-well plates. It is possible to derive from these libraries subsets of genes — for example, all G-protein-coupled receptors (GPCRs) or all putative druggable targets[30]. For screening, dsRNA material or dsRNA-expressing bacteria are transferred to 96-well plates containing *C. elegans* animals and phenotypes are scored over the following several days. The screening procedure itself is straightforward; the challenge lies in choosing the appropriate *C. elegans* disease model and in developing and validating the corresponding assay[28].

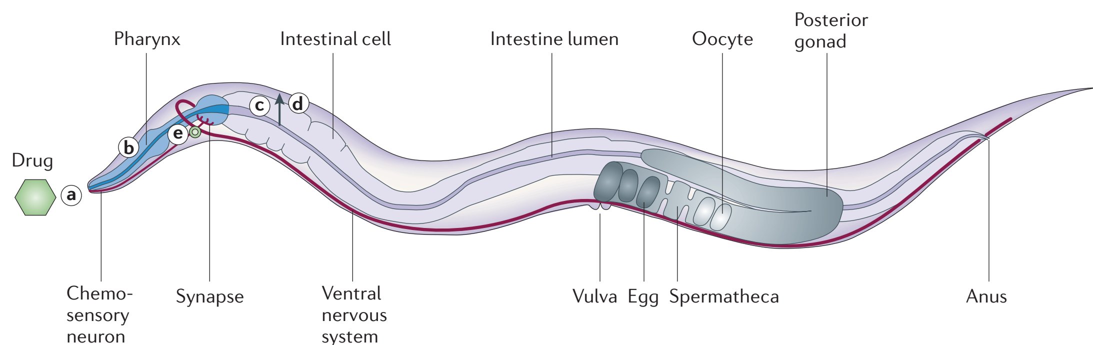

Figure 2 | **Drug entry route into *C. elegans*.** The simplified body plan of a *C. elegans* adult hermaphrodite. Drugs have been reported to enter *C. elegans* via three distinct routes: ingestion; uptake through the skin; and uptake via exposed sensory neuronal endings[93–95]. Uptake by ingestion is a relatively slow process (in the order of minutes) that involves several distinct steps. **a** | Chemosensory neurons help to choose appropriate food sources. Compounds with unpleasant smell or taste might be avoided and not taken up effectively. **b** | The drug is taken up by aspiration through the pharynx. The activity of this feeding organ depends on satiety or food availability and is controlled by various neurotransmitter systems. **c** | Once the drug has entered the intestinal lumen, it has to be absorbed by intestinal cells. The apical side facing the lumen is covered by microvilli, increasing the absorption surface. Indeed, the intestinal cells express many transport proteins such as P-glycoproteins and peptide transporters. **d** | From the body cavity the drug is distributed rapidly throughout the body. **e** | The drug reaches its target. In this example, the drug binds to receptors at a synapse in the central nervous system, evoking a behavioural response such as reduced feeding activity.

More than 15 genome-wide screens have been published in the past 6 years. Many of these screens were performed in the wild-type strain. Others took advantage of specific mutant backgrounds to develop more sensitive or selective assays. Below, we consider one such screen and how *C. elegans* can be used as a model for metabolic diseases.

Maintenance of metabolic health is crucial for preventing the risk of type 2 diabetes, hypertension, coronary heart disease, stroke and other diseases[31]. Obesity is the most common nutritional disorder and increases the likelihood of death by 20%. Only a handful of anti-obesity drugs have progressed to clinical trials so far. The opportunity to develop new drugs therefore remains compelling. There are two strategies for therapeutic intervention in obesity. First, food/energy uptake can be reduced by suppression of appetite or inhibition of dietary fat absorption. Second, fat storage can be reduced, for example, by decreasing triacylglycerol concentrations or by increasing metabolic activity to 'burn' fat. For example, sibutramine (Meridia; Abbott) acts on the central nervous system as an inhibitor of noradrenaline, serotonin (5-hydroxytryptamine; 5-HT) and dopamine reuptake to suppress appetite[32]. Orlistat (Xenical; Roche) blocks the activity of the gastrointestinal lipase, which reduces the processing of dietary fat and consequently reduces fat absorption[33].

The lifespan of *C. elegans* is also dependent on metabolic health. A high-calorie diet shortens lifespan, whereas caloric restriction can extend it[34]. Similarly, *C. elegans* mutants, in which feeding behaviour has been genetically reduced, enjoy a 50% increase in lifespan[35]. The body fat content of *C. elegans* is regulated by metabolic activity and feeding behaviour, which provides the opportunity to use *C. elegans* to identify genes that regulate feeding/satiety, nutritional uptake and fat metabolism.

Body fat content in *C. elegans* can be measured by Nile Red staining of tissue lipids followed by fluorescence detection. This read-out has been validated by correlating the Nile Red staining intensity with the actual body fat content as determined by gas chromatography in both wild-type worms and mutants with altered fat metabolism[36]. The Nile Red read-out is amenable to assay automation and has been used in a genome-wide RNAi screen for genes that, on RNAi-mediated knockdown, either increase or decrease the fat content of *C. elegans*. In this genome-wide campaign, several genetic *C. elegans* models were used. For example, *C. elegans* *daf-2* mutants with reduced insulin signalling accumulate high amounts of body fat. This model therefore emulates two human pathogeneses — reduced insulin signalling and high body fat. As expected, RNAi knockdown of genes such as the forkhead box O orthologue *daf-16* or the

Box 2 | Orthologues, orthologue clusters and inparalogues

Orthologues are genes in two species that have directly evolved from a single gene in the last common ancestor (for example, human and rat α-globin genes). Orthologues often have the same function in the two species. By contrast, paralogues are generated when gene duplication events lead to the multiplication of a single ancestral gene. Genes that diverged following gene duplication events (for example, human α- and β-globin) often have related but distinct functions. If a gene duplicated after two species diverged during evolution, then more than one orthologue exists in at least one of the two species; these are called inparalogues. Inparalogues have one-to-many or many-to-many relationships and are grouped into orthologue clusters[9]. Gene functions present in the original ancestor gene are often split up among inparalogues. For example, deficiencies in α-galactosidase A cause the human lysosomal storage disorder (LSD) Fabry's disease and deficiencies in the α-*N*-acetylgalactosaminidase cause the human LSD Schindler's disease. *Gana-1* is the *C. elegans* orthologue of both human genes α-galactosidase A and α-*N*-acetylgalactosaminidase[72]. Interestingly, *C. elegans* GANA-1 has dual enzymatic activity, combining both activities of the two human enzymes. Another example is provided by the BCL2 family of apoptosis regulators, which includes pro-survival genes such as *bcl-2*, and pro-apoptotic genes such as *bax*. In *C. elegans*, there is a single family member, CED-9, that is orthologous to both inparalogues BCL2 and BAX. Interestingly, CED-9 has both pro-survival and pro-apoptotic activities[73,74].

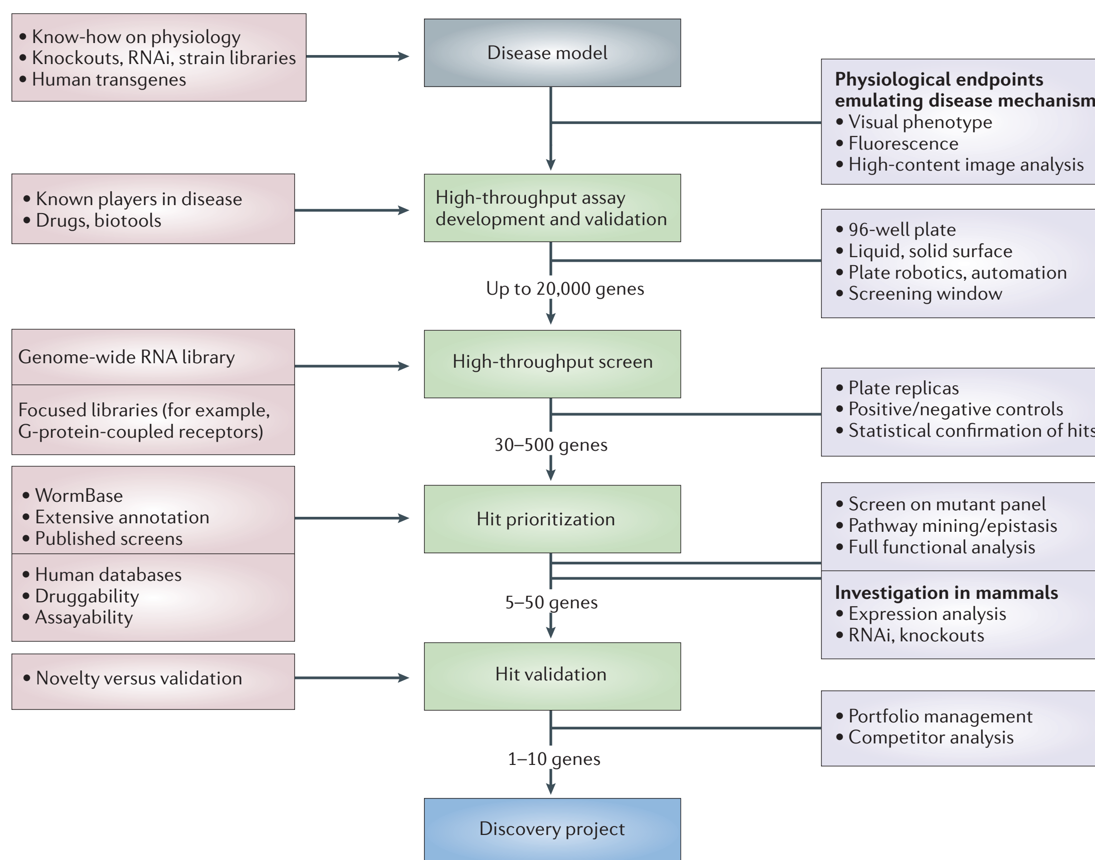

Figure 3 | **Overview of the RNA interference supported target identification process.** The first and crucial step is the design of a *C. elegans* disease model that emulates specific aspects of a disease mechanism such as reduced insulin signalling. The long-standing experience of the *C. elegans* field in large-scale mutagenesis screens combined with modern read-out technology allows developing medium- to high-throughput assays to perform genome-wide RNA interference (RNAi) screens. Assays need to be validated with a learning set of double stranded RNA (dsRNA) against genes that are known to modulate the disease pathology. Quality criteria need to be defined to ensure minimal day-to-day variation of the screening campaign. Screening results need to be statistically verified to avoid an excess of false positives and negatives. A wealth of public and proprietary information, databases and knowledge helps to prioritize hits. The cycle of hit prioritization and validation needs to satisfy two principal questions: what is the relevance of the human homologue as a potential target for disease treatment, and what is the probability of commercial success for a putative drug against the selected target? *C. elegans* can effectively support target prioritization because timelines to unravel gene function are short and costs are relatively low.

phosphatase and tensin homologue (PTEN) orthologue *daf-18*, reduced body fat back to wild-type level in this assay, due to restored insulin signalling. Furthermore, RNAi knockdowns of the orthologues of acetyl-coenzyme A carboxylase (ACC), fatty-acid synthase (FAS) and fatty-acid desaturases (SCD) reduced body fat in this assay as well. Another *C. elegans* model, used in this work, carried a mutation in the *C. elegans* orthologue of the mouse *tubby* gene. Similar to the obese *tubby* mice, *C. elegans tub-1* mutants had a 35% increase in fat content[36].

This assay was used for a genome-wide screen. Of the 16,757 genes tested, 417 genes were found to modulate fat content in *C. elegans*. The mammalian homologues of half of these *C. elegans* fat regulatory genes had not been previously implicated in regulating fat metabolism[36]. The power of *C. elegans* technology enabled this assay to be carried out genome-wide in wild-type worms and three genetic models. Furthermore, next to the actual screening data, information about viability and development has been generated for all genes. The combination of these results enables the classification of these genes into either candidate targets to control fat metabolism or targets to control feeding behaviour. Another advantage is that, unlike the list of genes obtained from transcriptional profiling studies, a genetic function for these genes has been already established; the reversion of the phenotype in the *C. elegans* disease model.

C. elegans toolbox, WormBase and proteomics

The central online platform for *C. elegans* information is WormBase. WormBase is an excellent 'one-stop shop' for *C. elegans* genetic and genomics resources (BOX 3).

WormBase, founded in 2000, is developed and curated by an international consortium of biologists and computer scientists. It has become a comprehensive resource for *C. elegans* biology, genomics and proteomics[37,38]. WormBase

Box 3 | **WormBase and other useful databases**

WormBase

The following information is contained directly in WormBase (http://www.wormbase.org):

- Gene summary with links to all available data generated for this gene, resources and literature links; blast searches; map positions; batch gene processing; and comparison with other organisms.
- Genomic sequences and an associated genome browser.
- Genome-wide RNA interference (RNAi): contains all published genome-wide RNAi campaigns in a searchable format.
- Genome-wide microarray: contains all published microarray expression data.
- WormBase also provides direct links to other external databases.

Textpresso

A full-text search engine of *C. elegans* literature (http://www.textpresso.org/).

ORFeome project

Provides all predicted open reading frames (ORFs) as Gateway Entry clones. *C. elegans* ORFeome version 3.1 contains 12,625 full-length protein-coding ORFs (http://worfdb.dfci.harvard.edu).

Worm Interactome

The WI5 map contains ~5,500 protein interactions (http://vidal.dfci.harvard.edu).

***C. elegans* SAGE data**

Serial analysis of gene expression of various developmental stages and FACS sorted cells and tissues (http://elegans.bcgsc.ca/home/sage.hmtl).

Genome-wide gene expression patterns

http://129.11.204.86:591/default.htm; http://elegans.bcgsc.ca/perl/eprofile/index

Structural genomics of *C. elegans*

High-throughput protein expression, crystallization and structure modelling database (http://sgce.cbse.uab.edu/index.php).

***Caenorhabditis* Genetic Center**

The most extensive public resource for *C. elegans* strains (http://biosci.umn.edu/CGC/CGChomepage.htm).

Wormatlas

A database of behavioural and structural anatomy of *C. elegans* (features, for example, the 'Slidabel worm') (http://www.wormatlas.org).

EST collection of Yuji Kohara

http://www.ddbj.nig.ac.jp/

***Caenorhabditis elegans* WWW Server**

Leon Avery provides links to many sites related to *C. elegans* research (http://elegans.swmed.edu/).

includes information on all genes, including a summary of protein function, NCBI KOGs (euKaryotic Orthologous Groups), gene models, genetic data, mutant and RNAi phenotypes, expression and interaction data, homology searches, literature and several data-mining tools.

From the many approaches used to investigate biological phenomena, we would like to draw special attention to the emerging efforts in *C. elegans*-based proteomics. In the post-genomics era, proteome-scale studies of protein structure, function and interaction are becoming increasingly important, both for basic research and therapeutic development. *C. elegans* is traditionally thought of as a classical genetic organism, but because the investigation of protein function is generally more complex than genome studies, *C. elegans* has the potential to pioneer genome-wide proteome analysis in the context of a whole animal. Recent publications suggest that this potential is already starting to be exploited[39].

The *C. elegans* ORFeome collection currently contains 65% of all predicted *C. elegans* open reading frames (ORFs) in Gateway entry vectors. This collection has been used for high-throughput expression of *C. elegans* proteins in bacteria[40]. For 4,854 of the 10,167 clones tested, protein expression was observed[41]. About 1,536 proteins were soluble and will be further processed for structure determination.

Another important proteomics approach is high-throughput yeast-two hybrid screening, because investigations of protein–protein interactions are crucial for a complete understanding of protein function. The current Worm Interactome map (WI5) contains 5,534 interactions derived from literature and high-throughput yeast-two hybrid approaches and connects about 15% of all *C. elegans* genes[42].

Chemical genetics and mode of drug action

Genetic analysis has been successfully used for the study of genes and proteins for nearly a century. In classical or forward genetics, the genome of an organism is randomly mutagenized. Mutants with the desired phenotype are collected and used to identify and characterize genes involved in the process of interest[43]. Chemical genetics is a special type of genetic screen that makes use of small-molecule compounds[44–46]. The principle is to screen for mutants that are either resistant or hypersensitive to the effect of a compound. These mutants carry genetic alterations either in the gene encoding the cognate target or in other genes acting in the same pathway. The following example illustrates the principle. Treatment of *C. elegans* with the antidepressant fluoxetine (Prozac; Eli

Chemical genetics
The investigation of the effect of a small molecule on an organism when a certain gene activity is increased or decreased; contrasts with chemical genomics studies, which look at the effect of a small-molecule compound on gene expression.

Cognate target
The protein to which a compound binds to mediate its physiological effect.

Box 4 | **Genetic screens in *C. elegans***

There are probably as many different genetic screens in *C. elegans* as there are geneticists. The following simplified procedure emphasizes some characteristic features of genetic screens in *C. elegans*[43]. A typical worm screen uses the mutagen ethylmethanesulphonate (EMS), which induces G/C to A/T transitions and small deletions. Hermaphrodites (P0) are incubated in 50 mM EMS for 4 hours — enough to inactivate 10–20 genes per genome. After treatment, the worms are distributed on Petri dishes and left to grow for two generations, resulting in homozygous mutants. The existence of hermaphrodites greatly facilitates *C. elegans* genetics because a strain can be maintained without setting up crosses. The progeny (F2) is screened for the desired phenotype. Isolated mutant strains can be conveniently preserved for long-term storage in liquid nitrogen. An F2 screen for recessive mutations can be completed within just 2 weeks. The high mutation frequency requires that the mutants are out-crossed to clean the genome of undesired mutations.

Genetic screens in *C. elegans* are efficient: screening of only 10,000–20,000 haploid genomes is sufficient to recover a few mutants per gene. Moreover, *C. elegans* screens can be up-scaled to screen millions of genomes. Such high-density screening campaigns allow the isolation of mutations even in very small regions of the genome, such as the binding site of a compound. The availability of detailed genetic and physical maps and the technologies to perform high-throughput gene mapping using single nucleotide polymorphisms make positional cloning in *C. elegans* fast and cost-effective[75–78].

Lilly) stimulates a number of behaviours, such as feeding, which are modulated by 5-HT signalling. Animals carrying a loss-of-function mutation in the *C. elegans* serotonin reuptake transporter (SERT) MOD-5 are resistant to fluoxetine action regarding these behaviours, identifying MOD-5 as a candidate target of fluoxetine[47,48]. Indeed, the MOD-5 homologue SERT is the major target of fluoxetine action in mammals.

Forward chemical genetics can be used to identify the mode of action of natural product molecules or lead compounds arising from phenotype screens. Identification of the target allows for the development of assays to enable lead optimization or the identification of further chemical hits and leads. Bristol-Myers Squibb Pharmaceuticals in collaboration with Exelixis used this approach to identify cellular targets of farnesyl

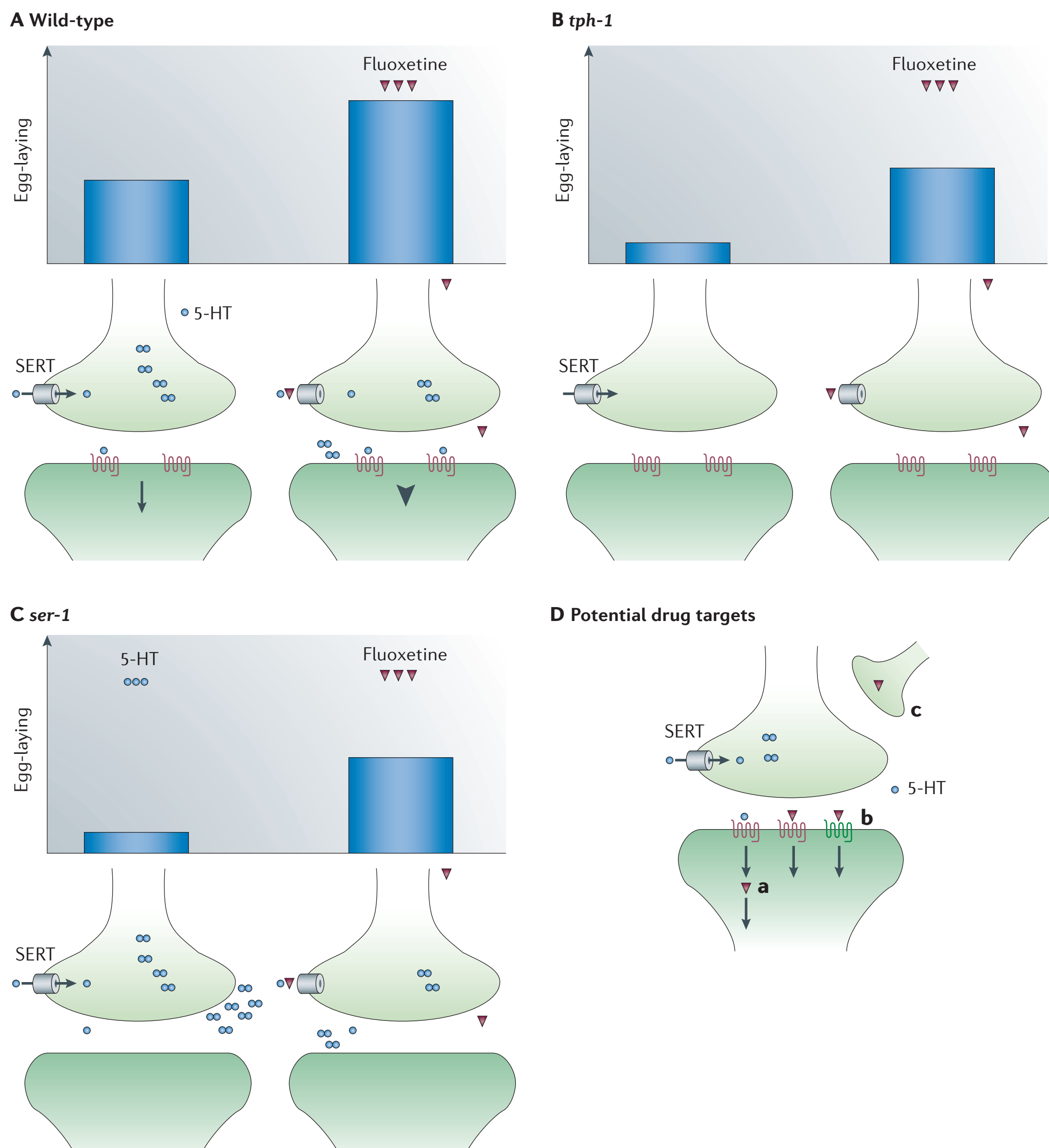

Figure 4 | **Chemical genetics as a tool for target site identification.** The graph (top half) shows the changes in egg-laying observed in an *in vivo C. elegans* experiment[53]. The cartoon (bottom half) illustrates the underlying mechanism. **A** | Fluoxetine blocks the serotonin-reuptake transporter (SERT) and thereby increases the concentration of serotonin (5-hydroxytryptamine; 5-HT) in the synaptic cleft. The resulting enhancement of serotonergic signalling stimulates egg-laying. **B** | In a genetic mutant, *tph-1*, that lacks tryptophan hydroxylase and therefore serotonin synthesis, inhibition of SERT by fluoxetine cannot increase the serotonin concentration. Surprisingly, egg-laying is nevertheless increased, suggesting additional interaction sites for fluoxetine. **C** | *C. elegans* animals lacking the gene encoding the 5-HT receptor (*ser-1*) cannot respond to exogenous applied serotonin but still respond to fluoxetine. **D** | Fluoxetine and other antidepressants could increase the synaptic transmission by a number of additional interactions: they can interact with downstream components in the postsynaptic cell (**a**), modulate activity of serotonin receptors or other receptors (**b**), or modulate presynaptic targets (**c**).

Box 5 | **High-throughput *C. elegans* egg-laying assay**

Egg-laying activity is a useful endpoint for studying processes related to the central nervous system and can be quantified simply by counting the number of eggs laid within a defined time interval. Egg-laying behaviour can also be measured in medium to high-throughput fashion, albeit indirectly, with the 'chitinase assay' developed by Pharmacia & Upjohn[79]. *C. elegans* embryos secrete chitinase to enable hatching out of the chitin-containing eggs. The total chitinase activity of a well therefore reflects the amount of hatching larvae, which is proportional to the number of eggs laid. Pfizer used this assay to screen more than 10,000 compounds in a *C. elegans* model of Alzheimer's disease[80].

transferase inhibitors (FTIs)[49]. FTIs block the post-translational modification of several proteins, including Ras, and have been developed as anticancer agents[50,51]. A number of these compounds have been shown to specifically revert the phenotype of Ras gain-of-function *C. elegans* mutants, suggesting that *C. elegans* could serve as a model for Ras-dependent cancer[52]. A set of these compounds also had pro-apoptotic activity against mammalian cells that do not carry activated Ras mutations, suggesting the presence of additional targets. To challenge the *C. elegans* model, these FTIs were tested for their pro-apoptotic activity in *C. elegans*. The *C. elegans* model responded to the FTIs with both the Ras-dependent as well as the pro-apoptotic phenotype, and so a sufficient degree of validation was achieved, at least for this compound class. In a second step, genetic and RNAi screens were performed to identify 74 genes that phenocopied FTI treatment (BOX 4). It is important to recognize that not all 74 genes can be targets of these FTIs; it is likely that some modulate the FTI pathway only indirectly, some affect a *C. elegans*-specific process, and others are involved in a general resistance mechanism. To avoid expending further effort on these false positives, they need to be filtered out. Indeed, prioritization of these genes led the authors to focus on the enzyme Rab geranyl-geranyl transferase (RabGGT) as a candidate target of FTIs in *C. elegans* animals. In a last step, this gene was validated in mammalian cancer cells by siRNA. Interestingly, RabGGT was found to be overexpressed in certain human tumour samples. In summary, this chemical genetics study both identified RabGGT as a novel target for cancer therapy and led to a better understanding of the mode of action of FTIs.

Chemical genetics has also been used to scrutinize the mode of action of antidepressants within the 5-HT pathway. Fluoxetine and imipramine are potent antidepressants that increase 5-HT at the synapse via inhibition of the SERT. Interestingly, the action of these two drugs cannot be explained solely by increasing 5-HT levels, which led to the suggestion that they act on additional targets. Indeed, both drugs bind to 5-HT receptors *in vitro*. However, it has not been resolved whether, and to what extent, these targets could contribute to the therapeutic action of the two drugs. To address this issue, Dempsey *et al.*[53] performed a detailed chemical genetics study of the effect of fluoxetine and imipramine on egg-laying in *C. elegans*. This behaviour is tightly controlled via various neurotransmitter systems such as 5-HT, acetylcholine and neuropeptides, to coordinate it with other behavioural activities of the animal.

Porsolt forced-swim test
Rats that are forced to swim in a cylinder from which they cannot escape will, after an initial period of vigorous activity, adopt a characteristic immobile posture that can be readily identified. Antidepressants will delay onset of immobility.

Exogenous 5-HT, fluoxetine and imipramine increase egg-laying mainly by increasing 5-HT levels at the synapse (FIG. 4). However, the following two experiments suggest that the effects on egg-laying cannot be explained with drug action on SERT alone.

First, the egg-laying defect of *tph-1* mutants, which lack endogenous 5-HT, could be restored not only by administration of exogenous 5-HT but also partially by fluoxetine and imipramine. These mutant animals should have been resistant to SERT inhibitors if these molecules exert their action solely through increasing 5-HT concentrations, because *tph-1* mutants completely lack 5-HT. Second, mutants lacking the 5-HT receptor homologue *ser-1* did not respond to exogenous 5-HT because the 5-HT signal is primarily mediated by the *ser-1*-dependent pathway. Again, fluoxetine and imipramine could increase egg-laying in *ser-1*-defective mutants, suggesting that these two drugs act on additional targets in the network controlling egg-laying. Dempsey *et al.* proposed that imipramine acts on another 5-HT receptor subtype (*ser-4*), whereas fluoxetine acts on a GPCR signalling pathway (FIG. 4). The implication of this study is that these drugs are even more potent than 5-HT itself because they not only increase 5-HT concentrations at the synapse, but also synergistically activate other pathways within the serotonergic network.

The key message of this study is that *C. elegans* genetics can be used to dissect the *in vivo* action of a drug, even if it modulates several targets. Indeed, *C. elegans* is an ideal system for bridging the gap between *in vitro* assays and mammalian models. *In vitro* assays provide biochemical information on the interactions between drug and target, but cannot predict which, and to what extent, these interactions will translate into *in vivo* activity. Mammalian animal models might define the best preclinical model for therapeutic intervention and provide evidence for *in vivo* efficacy, but they are too complex and tedious to dissect the mechanism of action. For example, the Porsolt forced-swim test predicts reasonably well antidepressant drug activity, but is not high-throughput. *C. elegans* has the benefit of providing *in vivo* drug activity in a whole animal setting, yet is simple enough to permit dissection of the mechanism of action. As the egg-laying assay can be performed in medium to high-throughput against a panel of mutants representing different aspects of the signalling network, such a *C. elegans* module could be integrated cost-effectively into the drug discovery pipeline (BOX 5).

Modelling human diseases in worms

A large number of human diseases have been investigated using *C. elegans* (TABLE 1). There are three general ways to generate a *C. elegans* disease model. First, knocking out (mutant) or knocking down (RNAi) the *C. elegans* homologue of the human disease gene in order to investigate the resulting phenotype; second, choosing a process in *C. elegans* that reproduces certain molecular or cellular aspects of the disease mechanism; and third, expressing the human gene to induce a disease-related phenotype in *C. elegans*. This approach is particularly useful if the gene product interferes with normal cell function.

All three approaches need stringent validation to gain confidence in the model and to better define which specific aspects of the human disease can be modelled. In particular, the first approach is prone to disappointments if the investigator expects the worm mutant phenotype to fully recreate the descriptive disease

Table 1 | **C. elegans disease models**

Disease	Pathway or genes (*C. elegans* orthologue)	Model and method of validation
Metabolic syndrome		
Diabetes and obesity	Insulin/AKT (*daf-2*, *ins-1*, *akt-1*)	Dauer formation model for insulin signalling was validated with rescue of heterologously expressed human insulin[96]. Model used by Devgen and Exelixis for genome-wide RNAi screens[97]. Fat metabolism model was validated with *daf-16* (FOXO) and *fat-7* (SCD)[36].
Ageing	Caloric restriction; oxidative stress, for example, manganese superoxide dismutase (*sod-3*); insulin receptor/ PI3K (*daf-2*, *age-1*)	Longevity model: vitamin E, caloric restriction and reduced insulin signalling prolonged life in *C. elegans*[98–100]. Mammalian validation: studies in insulin receptor mutant mice carrying the homologous mutation of *daf-2* mutants demonstrated that reduced insulin signalling increases resistance to oxidative stress[101–103].
Oncology		
Cancer	EGF/RAS (*let-23*, *let-60*) DNA damage (for example, p53 (*cep-1*), BRCA1 (*brc-1*))	Vulva development model for cancer biology[104,105]: this model was validated through farnesyl transferase inhibitors, which were capable of restoring excessive vulva formation in RAS over-activated *C. elegans*[52]. Ionization-induced apoptosis model: validated by demonstrating the conserved interactions of oncoprotein iASPP and c-Abl with p53[106,107]. BRCA1 pathway mining[108].
Neurodegeneration		
Alzheimer's disease	Presenilin (*sel-12*) Aβ	Genetic egg-laying model and neuronal model: both models were validated by demonstration that human presenilin can substitute for endogenous *C. elegans* presenilin[65,66]. Transgenic model: expression of human β-amyloid (Aβ) peptide in *C. elegans* induced formation of amyloid deposits, which in turn caused paralysis[109–112].
Parkinson's disease	α-Synuclein (no orthologue)	Transgenic model: overexpression of human α-synuclein caused neuronal and dendritic loss in *C. elegans*[113]. Pharmacological model: 6-OHDA (6-hydroxydopamine) a precursor of 1-methyl-4-phenylpyridinium (MPP) induced selective dopamine neuronal death similar to vertebrate models[114].
Huntington's disease	Huntingtin (no orthologue); polyglutamine (polyQ) aggregation	Transgenic model: heterologous expression of a Huntingtin fragment with expanded polyQ in *C. elegans* caused specific polyQ-length-dependent neuronal malfunction and the formation of aggregates[115]. Genome-wide RNAi screen for regulators of polyQ aggregation using fluorescence read-out[116,117].
Neurobiology		
Depression	Serotonin	Behavioural changes after treatment with antidepressants (see text).
Pain, neuronal regeneration	Regeneration	Phenotype model: femtosecond laser operated GFP-labelled axons of *C. elegans* regenerated within 24 hours[118].
Genetic diseases		
ADPKD	PKD1 (*lov-1*), PKD2 (*pkd-2*)	Disease gene knockdown model: *C. elegans* mating behaviour is altered in PKD-deficient ciliated neurons. This provided the first evidence of the function of the polycystin as fluid-sensors in ciliated cells[55,57].
Muscular dystrophy	Dystrophin (*dys-1*), emerin (*emr-1*)	Disease gene knockdown model: *C. elegans* animals with mutations in *dys-1* develop a muscle degeneration phenotype. This can be partially treated with prednisone[119]. The specific locomotion phenotype has been used to identify *snf-6*, a sodium-dependent neurotransmitter transporter, which has a role in the disruption of the dystrophin–glycoprotein complex (DGC)[120].
Ionchannelopathies	Ion channels	Overexpression of human channels in the pharynx to change *C. elegans* action potential. Devgen is using this approach to perform high-throughput compound screening (30,000 data points per week)[48].
Innate immunity	*Bacillus thuringiensis* toxin (Bt toxin) Host–pathogen interaction	*C. elegans* is killed by Bt toxin, a major pesticidal agent. A number of *C. elegans* resistance screens have identified resistance mechanisms[121]. *C. elegans–Serratia marcescens* host–pathogen model was used to identify virulence factors. Model confirmed by identifying bacterial clones with reduced virulescence in *C. elegans* as well as in cells[122].

ADPKD, autosomal dominant polycystic kidney disease; BRCA1, breast cancer gene 1 ; EGF, epidermal growth factor; PI3K, phosphatidylinositol 3-kinase; PKD, polycystin; RNAi, RNA interference.

pathology. A compelling example is the use of *C. elegans* to model autosomal dominant polycystic kidney disease (ADPKD). The primary pathology of ADPKD is the development of cystic kidneys, which gradually progress to end-stage renal failure[54]. As *C. elegans* has only a very rudimentary kidney-like organ (consisting of a single secretory cell), it is at first glance debatable whether it would make any sense to use *C. elegans* as a model of ADPKD. Gene knockdown and knockout of *lov-1* and *pkd-2*, the *C. elegans* orthologues of the human disease genes *PKD1* and *PKD2*, did not induce any kind of cystic structures in *C. elegans* but instead caused aberrant mating behaviour[55,56]. Although the *C. elegans* phenotype was counter intuitive, further worm work eventually determined the proper function of the polycystins. Detailed analysis of the mating behaviour revealed that LOV-1 and PKD-2 have a crucial signalling role in the ciliated endings of male-specific sensory neurons, which led to the hypothesis that the human polycystins are required for ciliated cells to sense pressure or fluid flow[57]. Studies in mice confirmed that the polycystins do indeed mediate fluid flow in the primary cilia of renal epithelial cells[58,59]. Although this example might be an extreme case, it shows the power of *C. elegans* genetics to deconvolute even complex disease pathology.

The third approach, the development of transgenic models, has been intensively used for investigating neurodegenerative disorders such as Alzheimer's disease, Parkinson's disease, Huntington's disease and tauopathies. In some of these cases, the underlying genes — for example, Huntingtin, the human disease gene that is associated with Huntington's disease — do not have readily recognizable orthologues in *C. elegans*. Nevertheless, conserved responses or interactions can often be detected. For example, yeast-two hybrid screening of the *C. elegans* proteome with human Huntingtin identified homologues of a number of known Huntingtin-interacting proteins such as CA150 or CRE[60]. *C. elegans* models have helped to elucidate disease pathogenesis, suggested new approaches for therapeutic interventions, and provided whole-animal assays amenable to high-throughput target and drug screening.

C. elegans is also emerging as an attractive genetic model for the study of host–pathogen interactions. *C. elegans* is killed by many pathogens such as *Pseudomonas aeruginosa*, *Serratia marcescens*, *Salmonella enterica*, *Staphylococcus aureus* and *Streptococcus pneumoniae*. Many virulence mechanisms used by pathogens to cause disease in humans have been shown to be important for disease in *C. elegans*, including persistent infection of the intestine, colonization with biofilm formation on the worm cuticle, and killing via toxins such as Botulinum toxin, hydrogen cyanide or hydrogen peroxide[61]. Similarly, components of the host-encoded defence system are evolutionarily conserved[62].

Microbial virulence factors such as the quorum-sensing systems or the Gram-negative type III secretion system are important for worm killing[63]. The high degree of overlap between the virulence factors required for pathogenesis in nematodes and humans validates the use of *C. elegans* as a surrogate host. Pathogenic strains can therefore be screened on *C. elegans* to identify virulence factors or, vice versa, *C. elegans* mutants can be screened against a pathogen to better understand mechanisms of pathogenesis and host response.

Conclusions

Many biological processes are conserved between humans and *C. elegans* to such an extent that *C. elegans* data are in many cases predictive for drug–target interaction and target validation. For a number of human diseases, *C. elegans* investigations have already fostered a better understanding of the underlying mechanism. Despite its many advantages, the *C. elegans* system of course also has intrinsic limitations that could limit its application in drug discovery. For example, some molecular pathways simply do not exist in the worm, and therefore cannot be studied there. In addition, *C. elegans* models usually do not recreate the complete pathophysiology of the human disease. Therefore, defining the purpose and potential restrictions of the model is crucial for success. Emphasis must be laid on validating the *C. elegans* disease models to ensure congruence between the model and specific aspects of the human pathology. When these conditions are met, however, the value of *C. elegans* readily becomes apparent. Indeed, *C. elegans* bridges the gap that intrinsically exists between *in vitro* and *in vivo* approaches because it allows a high-throughput, reductionist approach while at the same time providing physiologically relevant data derived from a whole-animal setting. Judiciously inserted, *C. elegans* models can be a useful addition to many a drug discovery pipeline.

1. Brenner, S. The genetics of *Caenorhabditis elegans*. *Genetics* **77**, 71–94 (1974).
2. Riddle, D., Blumenthal, T., Meyer, B. & Priess, J. C. elegans *II* (Cold Spring Harbor Laboratory, New York, 1997).
3. The *C. elegans* Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**, 2012–2018 (1998).
4. White, J. in *The Nematode* Caenorhabditis elegans (ed. Wood, W. B.) 81–122 (Cold Spring Harbor Laboratory, New York, 1988).
5. Sonnhammer, E. L. & Durbin, R. Analysis of protein domain families in *Caenorhabditis elegans*. *Genomics* **46**, 200–216 (1997).
6. Lai, C. H., Chou, C. Y., Ch'ang, L. Y., Liu, C. S. & Lin, W. Identification of novel human genes evolutionarily conserved in *Caenorhabditis elegans* by comparative proteomics. *Genome Res.* **10**, 703–713 (2000).
7. Kuwabara, P. E. & O'Neil, N. The use of functional genomics in *C. elegans* for studying human development and disease. *J. Inherit. Metab Dis.* **24**, 127–138 (2001).
8. Harris, T. W. *et al*. WormBase: a multi-species resource for nematode biology and genomics. *Nucleic Acids Res.* **32**, D411–D417 (2004).
9. Remm, M., Storm, C. E. & Sonnhammer, E. L. Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J. Mol. Biol.* **314**, 1041–1052 (2001).
10. O'Brien, K. P., Westerlund, I. & Sonnhammer, E. L. OrthoDisease: a database of human disease orthologs. *Hum. Mutat.* **24**, 112–119 (2004).
11. Schuske, K., Beg, A. A. & Jorgensen, E. M. The GABA nervous system in *C. elegans*. *Trends Neurosci.* **27**, 407–414 (2004).
12. McIntire, S. L., Reimer, R. J., Schuske, K., Edwards, R. H. & Jorgensen, E. M. Identification and characterization of the vesicular GABA transporter. *Nature* **389**, 870–876 (1997).
13. Dumoulin, A. *et al*. Presence of the vesicular inhibitory amino acid transporter in GABAergic and glycinergic synaptic terminal boutons. *J. Cell Sci.* **112**, 811–823 (1999).

14. Drews, J. Drug discovery: a historical perspective. *Science* **287**, 1960–1964 (2000).
15. Hopkins, A. L. & Groom, C. R. The druggable genome. *Nature Rev. Drug Discov.* **1**, 727–730 (2002).
16. Fire, A. *et al*. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811 (1998).
17. Fraser, A. G. *et al*. Functional genomic analysis of *C. elegans* chromosome 1 by systematic RNA interference. *Nature* **408**, 325–330 (2000).
18. Gonczy, P. *et al*. Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome 3. *Nature* **408**, 331–336 (2000).
19. Mello, C. C. & Conte, D. Jr. Revealing the world of RNA interference. *Nature* **431**, 338–342 (2004).
20. Plaetinck, G., Platteeuw, C., Mortier, K. & Bogaert, T. Characterization of gene function using double stranded RNA inhibition. WO Patent 00/01846 (1999).
21. Timmons, L. & Fire, A. Specific interference by ingested dsRNA. *Nature* **395**, 854 (1998).
22. Tabara, H., Grishok, A. & Mello, C. C. RNAi in *C. elegans*: soaking in the genome sequence. *Science* **282**, 430–431 (1998).
23. Kamath, R. S. & Ahringer, J. Genome-wide RNAi screening in *Caenorhabditis elegans*. *Methods* **30**, 313–321 (2003).
24. Winston, W. M., Molodowitch, C. & Hunter, C. P. Systemic RNAi in *C. elegans* requires the putative transmembrane protein SID-1. *Science* **295**, 2456–2459 (2002).
25. Elbashir, S. M. *et al*. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494–498 (2001).
26. Yang, S., Tutton, S., Pierce, E. & Yoon, K. Specific double-stranded RNA interference in undifferentiated mouse embryonic stem cells. *Mol. Cell Biol.* **21**, 7807–7816 (2001).
27. Jackson, A. L. *et al*. Expression profiling reveals off-target gene regulation by RNAi. *Nature Biotechnol.* **21**, 635–637 (2003).
28. Sonnichsen, B. *et al*. Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. *Nature* **434**, 462–469 (2005).
29. Kamath, R. S., Martinez-Campos, M., Zipperlen, P., Fraser, A. G. & Ahringer, J. Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*. *Genome Biol.* **2**, (2001).
30. Keating, C. D. *et al*. Whole-genome analysis of 60 G protein-coupled receptors in *Caenorhabditis elegans* by gene knockout with RNAi. *Curr. Biol.* **13**, 1715–1720 (2003).
31. Shi, Y. & Burn, P. Lipid metabolic enzymes: emerging drug targets for the treatment of obesity. *Nature Rev. Drug Discov.* **3**, 695–710 (2004).
32. Filippatos, T. D., Kiortsis, D. N., Liberopoulos, E. N., Mikhailidis, D. P. & Elisaf, M. S. A review of the metabolic effects of sibutramine. *Curr. Med. Res. Opin.* **21**, 457–468 (2005).
33. Curran, M. P. & Scott, L. J. Orlistat: a review of its use in the management of patients with obesity. *Drugs* **64**, 2845–2864 (2004).
34. Klass, M. R. Aging in the nematode *Caenorhabditis elegans*: major biological and environmental factors influencing life span. *Mech. Ageing Dev.* **6**, 413–429 (1977).
35. Lakowski, B. & Hekimi, S. The genetics of caloric restriction in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **95**, 13091–13096 (1998).
36. Ashrafi, K. *et al*. Genome-wide RNAi analysis of *Caenorhabditis elegans* fat regulatory genes. *Nature* **421**, 268–272 (2003).
Excellent example of how to perform an RNAi screen in *C. elegans* on a genome-wide scale to identify targets for treatment of obesity and diabetes and subsequent prioritization of targets by parallel screening in several *C. elegans* metabolic disease models.
37. Stein, L., Sternberg, P., Durbin, R., Thierry-Mieg, J. & Spieth, J. WormBase: network access to the genome and biology of *Caenorhabditis elegans*. *Nucleic Acids Res.* **29**, 82–86 (2001).
38. Chen, N. *et al*. WormBase: a comprehensive data resource for *Caenorhabditis* biology and genomics. *Nucleic Acids Res.* **33**, D383–D389 (2005).
39. Duchaine, T. F. *et al*. Functional proteomics reveals the biochemical niche of *C. elegans* DCR-1 in multiple small-RNA-mediated pathways. *Cell* **124**, 343–354 (2006).
40. Reboul, J. *et al*. *C. elegans* ORFeome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression. *Nature Genet.* **34**, 35–41 (2003).
41. Luan, C. H. *et al*. High-throughput expression of *C. elegans* proteins. *Genome Res.* **14**, 2102–2110 (2004).
42. Li, S. *et al*. A map of the interactome network of the metazoan *C. elegans*. *Science* **303**, 540–543 (2004).
43. Jorgensen, E. M. & Mango, S. E. The art and design of genetic screens: *Caenorhabditis elegans*. *Nature Rev. Genet.* **3**, 356–369 (2002).
44. Jones, A. K., Buckingham, S. D. & Sattelle, D. B. Chemistry-to-gene screens in *Caenorhabditis elegans*. *Nature Rev. Drug Discov.* **4**, 321–330 (2005).
45. Zheng, X. F. & Chan, T. F. Chemical genomics: a systematic approach in biological research and drug discovery. *Curr. Issues Mol. Biol.* **4**, 33–43 (2002).
46. Alaoui-Ismaili, M. H., Lomedico, P. T. & Jindal, S. Chemical genomics: discovery of disease genes and drugs. *Drug Discov. Today* **7**, 292–294 (2002).
47. Ranganathan, R., Sawin, E. R., Trent, C. & Horvitz, H. R. Mutations in the *Caenorhabditis elegans* serotonin reuptake transporter MOD-5 reveal serotonin-dependent and -independent activities of fluoxetine. *J. Neurosci.* **21**, 5871–5884 (2001).
48. Kaletta, T., Butler, L. & Bogaert, T. in *Model Organisms in Drug Discovery* (eds Carroll, P. M. & Fitzgerald, K.) 41–79 (John Wiley & Sons, 2003).
49. Lackner, M. R. *et al*. Chemical genetics identifies Rab geranylgeranyl transferase as an apoptotic target of farnesyl transferase inhibitors. *Cancer Cell* **7**, 325–336 (2005).
The authors describe how to build and validate a *C. elegans* assay for the investigation of multiple *in vivo* modes of action of farnesyl transferase inhibitors (FTIs), to identify and prioritize targets of FTIs and to validate these novel targets in mammalian systems.
50. Karp, J. E. *et al*. Current status of clinical trials of farnesyltransferase inhibitors. *Curr. Opin. Oncol.* **13**, 470–476 (2001).
51. O'Regan, R. M. & Khuri, F. R. Farnesyl transferase inhibitors: the next targeted therapies for breast cancer? *Endocr. Relat. Cancer* **11**, 191–205 (2004).
52. Hara, M. & Han, M. Ras farnesyltransferase inhibitors suppress the phenotype resulting from an activated ras mutation in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **92**, 3333–3337 (1995).
53. Dempsey, C. M., Mackenzie, S. M., Gargus, A., Blanco, G. & Sze, J. Y. Serotonin (5HT), fluoxetine, imipramine and dopamine target distinct 5HT receptor signaling to modulate *Caenorhabditis elegans* egg-laying behavior. *Genetics* **169**, 1425–1436 (2005).
Good example of the conservation of human and *C. elegans* pharmacology at the cellular level.
54. Delmas, P. Polycystins: from mechanosensation to gene regulation. *Cell* **118**, 145–148 (2004).
55. Barr, M. M. & Sternberg, P. W. A polycystic kidney-disease gene homologue required for male mating behaviour in *C. elegans*. *Nature* **401**, 386–389 (1999).
56. Barr, M. M. *et al*. 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).
57. Kaletta, T. *et al*. Towards understanding the polycystins. *Nephron Exp. Nephrol.* **93**, e9–e17 (2003).
58. Nauli, S. M. *et al*. Polycystins 1 and 2 mediate mechanosensation in the primary cilium of kidney cells. *Nature Genet.* **33**, 129–137 (2003).
59. Lipton, J. Mating worms and the cystic kidney: *Caenorhabditis elegans* as a model for renal disease. *Pediatr. Nephrol.* **20**, 1531–1536 (2005).
60. Parker, J. A., Holbert, S., Lambert, E., Abderrahmane, S. & Neri, C. Genetic and pharmacological suppression of polyglutamine-dependent neuronal dysfunction in *Caenorhabditis elegans*. *J. Mol. Neurosci.* **23**, 61–68 (2004).
61. Sifri, C. D., Begun, J. & Ausubel, F. M. The worm has turned — microbial virulence modeled in *Caenorhabditis elegans*. *Trends Microbiol.* **13**, 119–127 (2005).
62. Kurz, C. L. & Ewbank, J. J. *Caenorhabditis elegans*: an emerging genetic model for the study of innate immunity. *Nature Rev. Genet.* **4**, 380–390 (2003).
63. Tenor, J. L., McCormick, B. A., Ausubel, F. M. & Aballay, A. *Caenorhabditis elegans*-based screen identifies *Salmonella* virulence factors required for conserved host–pathogen interactions. *Curr. Biol.* **14**, 1018–1024 (2004).
64. Sundaram, M. & Greenwald, I. Suppressors of a lin-12 hypomorph define genes that interact with both lin-12 and glp-1 in *Caenorhabditis elegans*. *Genetics* **135**, 765–783 (1993).
65. Levitan, D. & Greenwald, I. Facilitation of lin-12-mediated signalling by *sel-12*, a *Caenorhabditis elegans* S182 Alzheimer's disease gene. *Nature* **377**, 351–354 (1995).
66. Sherrington, R. *et al*. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature* **375**, 754–760 (1995).
67. Wittenburg, N. *et al*. Presenilin is required for proper morphology and function of neurons in *C. elegans*. *Nature* **406**, 306–309 (2000).
68. Levitan, D. *et al*. Assessment of normal and mutant human presenilin function in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **93**, 14940–14944 (1996).
69. De Strooper, B. *et al*. A presenilin-1-dependent γ-secretase-like protease mediates release of Notch intracellular domain. *Nature* **398**, 518–522 (1999).
70. Ogg, S. *et al*. The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in *C. elegans*. *Nature* **389**, 994–999 (1997).
71. Nakae, J. *et al*. Regulation of insulin action and pancreatic β-cell function by mutated alleles of the gene encoding forkhead transcription factor Foxo1. *Nature Genet.* **32**, 245–253 (2002).
72. Hujova, J. *et al*. Characterization of *gana-1*, a *Caenorhabditis elegans* gene encoding a single ortholog of vertebrate α-galactosidase and α-*N*-acetylgalactosaminidase. *BMC Cell Biol.* **6**, 5 (2005).
An interesting example of orthology clusters with between one and many relationships.
73. Jagasia, R., Grote, P., Westermann, B. & Conradt, B. DRP-1-mediated mitochondrial fragmentation during EGL-1-induced cell death in *C. elegans*. *Nature* **433**, 754–760 (2005).
74. Hengartner, M. O. & Horvitz, H. R. Activation of *C. elegans* cell death protein CED-9 by an amino-acid substitution in a domain conserved in Bcl-2. *Nature* **369**, 318–320 (1994).
75. Jakubowski, J. & Kornfeld, K. A local, high-density, single-nucleotide polymorphism map used to clone *Caenorhabditis elegans* cdf-1. *Genetics* **153**, 743–752 (1999).
76. Wicks, S. R., Yeh, R. T., Gish, W. R., Waterston, R. H. & Plasterk, R. H. Rapid gene mapping in *Caenorhabditis elegans* using a high density polymorphism map. *Nature Genet.* **28**, 160–164 (2001).
77. Swan, K. A. *et al*. High-throughput gene mapping in *Caenorhabditis elegans*. *Genome Res.* **12**, 1100–1105 (2002).
78. Zipperlen, P. *et al*. A universal method for automated gene mapping. *Genome Biol.* **6**, R19 (2005).
79. Gurney, M. E., Geary, T., Ellerbrbrock, B. R. & Thomas, E. M. A nematode drug screen for modulators of mammalian disorders. WO Patent 00/73493 (2000).
80. Ellerbrock, B. R., Coscarelli, E. M., Gurney, M. E. & Geary, T. G. Screening for presenilin inhibitors using the free-living nematode, *Caenorhabditis elegans*. *J. Biomol. Screen.* **9**, 147–152 (2004).
81. Desai, C., Garriga, G., McIntire, S. L. & Horvitz, H. R. A genetic pathway for the development of the *Caenorhabditis elegans* HSN motor neurons. *Nature* **336**, 638–646 (1988).
82. Desai, C. & Horvitz, H. R. *Caenorhabditis elegans* mutants defective in the functioning of the motor neurons responsible for egg laying. *Genetics* **121**, 703–721 (1989).
83. Weinshenker, D., Garriga, G. & Thomas, J. H. Genetic and pharmacological analysis of neurotransmitters controlling egg laying in *C. elegans*. *J. Neurosci.* **15**, 6975–6985 (1995).
84. Kim, J. *et al*. Genes affecting the activity of nicotinic receptors involved in *Caenorhabditis elegans* egg-laying behavior. *Genetics* **157**, 1599–1610 (2001).
85. Sze, J. Y., Victor, M., Loer, C., Shi, Y. & Ruvkun, G. Food and metabolic signalling defects in a *Caenorhabditis elegans* serotonin-synthesis mutant. *Nature* **403**, 560–564 (2000).
86. Loer, C. M. & Kenyon, C. J. Serotonin-deficient mutants and male mating behavior in the nematode *Caenorhabditis elegans*. *J. Neurosci.* **13**, 5407–5417 (1993).
87. Weimer, R. M. & Jorgensen, E. M. Controversies in synaptic vesicle exocytosis. *J. Cell Sci.* **116**, 3661–3666 (2003).

88. Hosono, R. & Kamiya, Y. Additional genes which result in an elevation of acetylcholine levels by mutations in *Caenorhabditis elegans*. *Neurosci. Lett.* **128**, 243–244 (1991).
89. Hosono, R. *et al*. The unc-18 gene encodes a novel protein affecting the kinetics of acetylcholine metabolism in the nematode *Caenorhabditis elegans*. *J. Neurochem.* **58**, 1517–1525 (1992).
90. White, J. G., Southgate, E., Thomas, J. N. & Brenner, S. The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Phil. Trans. R. Soc. London B* **314**, 1–340 (1986).
91. Lee, Y. S. *et al*. Characterization of GAR-2, a novel G protein-linked acetylcholine receptor from *Caenorhabditis elegans*. *J. Neurochem.* **75**, 1800–1809 (2000).
92. van der Linden, A. M., Simmer, F., Cuppen, E. & Plasterk, R. H. The G-protein β-subunit GPB-2 in *Caenorhabditis elegans* regulates the $G_o\alpha$–$G_q\alpha$ signaling network through interactions with the regulator of G-protein signaling proteins EGL-10 and EAT-16. *Genetics* **158**, 221–235 (2001).
93. Perkins, L. A., Hedgecock, E. M., Thomson, J. N. & Culotti, J. G. Mutant sensory cilia in the nematode *Caenorhabditis elegans*. *Dev. Biol.* **117**, 456–487 (1986).
94. Smith, H. & Campbell, W. C. Effect of ivermectin on *Caenorhabditis elegans* larvae previously exposed to alcoholic immobilization. *J. Parasitol.* **82**, 187–188 (1996).
95. Choy, R. K., Kemner, J. M. & Thomas, J. H. Fluoxetine-resistance genes in *Caenorhabditis elegans* function in the intestine and may act in drug transport. *Genetics* **172**, 885–892 (2006).
96. Pierce, S. B. *et al*. Regulation of DAF-2 receptor signaling by human insulin and ins-1, a member of the unusually large and diverse *C. elegans* insulin gene family. *Genes Dev.* **15**, 672–686 (2001).
97. Feichtinger, R. & Bogaert, T. Compound screens relating to insulin deficiency or insulin resistance. WO Patent 01/93669 (2001).
98. Harrington, L. A. & Harley, C. B. Effect of vitamin E on lifespan and reproduction in *Caenorhabditis elegans*. *Mech. Ageing Dev.* **43**, 71–78 (1988).
99. Kimura, K. D., Tissenbaum, H. A., Liu, Y. & Ruvkun, G. daf-2, an insulin receptor-like gene that regulates longevity and diapause in *Caenorhabditis elegans*. *Science* **277**, 942–946 (1997).
100. Tissenbaum, H. A. & Guarente, L. Model organisms as a guide to mammalian aging. *Dev. Cell* **2**, 9–19 (2002).
101. Houthoofd, K. *et al*. DAF-2 pathway mutations and food restriction in aging *Caenorhabditis elegans* differentially affect metabolism. *Neurobiol. Aging* **26**, 689–696 (2005).
102. Holzenberger, M. *et al*. IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice. *Nature* **421**, 182–187 (2003).
103. Baba, T. *et al*. Estrogen, insulin, and dietary signals cooperatively regulate longevity signals to enhance resistance to oxidative stress in mice. *J. Biol. Chem.* **280**, 16417–16426 (2005).
104. Sternberg, P. W. & Han, M. Genetics of RAS signaling in *C. elegans*. *Trends Genet.* **14**, 466–472 (1998).
105. Fisher, J., Piterman, N., Hubbard, E. J., Stern, M. J. & Harel, D. Computational insights into *Caenorhabditis elegans* vulval development. *Proc. Natl Acad. Sci. USA* **102**, 1951–1956 (2005).
106. Bergamaschi, D. *et al*. iASPP oncoprotein is a key inhibitor of p53 conserved from worm to human. *Nature Genet.* **33**, 162–167 (2003).
107. Deng, X. *et al*. *Caenorhabditis elegans* ABL-1 antagonizes p53-mediated germline apoptosis after ionizing irradiation. *Nature Genet.* **36**, 906–912 (2004).
108. Boulton, S. J. *et al*. BRCA1/BARD1 orthologs required for DNA repair in *Caenorhabditis elegans*. *Curr. Biol.* **14**, 33–39 (2004).
109. Link, C. D. Transgenic invertebrate models of age-associated neurodegenerative diseases. *Mech. Ageing Dev.* **122**, 1639–1649 (2001).
110. Link, C. D. Expression of human β-amyloid peptide in transgenic *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **92**, 9368–9372 (1995).
111. Westlund, B., Stilwell, G. & Sluder, A. Invertebrate disease models in neurotherapeutic discovery. *Curr. Opin. Drug Discov. Devel.* **7**, 169–178 (2004).
112. Link, C. D. Transgenic invertebrate models of age-associated neurodegenerative diseases. *Mech. Ageing Dev.* **122**, 1639–1649 (2001).
113. Lakso, M. *et al*. Dopaminergic neuronal loss and motor deficits in *Caenorhabditis elegans* overexpressing human α-synuclein. *J. Neurochem.* **86**, 165–172 (2003).
114. Nass, R., Hall, D. H., Miller, D. M., III & Blakely, R. D. Neurotoxin-induced degeneration of dopamine neurons in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **99**, 3264–3269 (2002).
115. Faber, P. W., Alter, J. R., MacDonald, M. E. & Hart, A. C. Polyglutamine-mediated dysfunction and apoptotic death of a *Caenorhabditis elegans* sensory neuron. *Proc. Natl Acad. Sci. USA* **96**, 179–184 (1999).
116. Nollen, E. A. *et al*. Genome-wide RNA interference screen identifies previously undescribed regulators of polyglutamine aggregation. *Proc. Natl Acad. Sci. USA* **101**, 6403–6408 (2004).
117. Buckingham, S. D., Esmaeili, B., Wood, M. & Sattelle, D. B. RNA interference: from model organisms towards therapy for neural and neuromuscular disorders. *Hum. Mol. Genet.* **13**, R275–R288 (2004).
118. Yanik, M. F. *et al*. Neurosurgery: functional regeneration after laser axotomy. *Nature* **432**, 822 (2004).
119. Gaud, A. *et al*. Prednisone reduces muscle degeneration in dystrophin-deficient *Caenorhabditis elegans*. *Neuromuscul. Disord.* **14**, 365–370 (2004).
120. Kim, H., Rogers, M. J., Richmond, J. E. & McIntire, S. L. SNF-6 is an acetylcholine transporter interacting with the dystrophin complex in *Caenorhabditis elegans*. *Nature* **430**, 891–896 (2004).
121. Griffitts, J. S. *et al*. Glycolipids as receptors for *Bacillus thuringiensis* crystal toxin. *Science* **307**, 922–925 (2005).
122. Kurz, C. L. *et al*. Virulence factors of the human opportunistic pathogen *Serratia marcescens* identified by *in vivo* screening. *EMBO J.* **22**, 1451–1460 (2003).

Competing interests statement
The authors declare competing financial interests: see Web version for details.

DATABASES
The following terms in this article are linked online to:
Entrez Gene:
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene
daf-2 | *daf-16* | *daf-18* | *GABRA6* | *lov-1* | MOD-5 | *PKD1* | *PKD2* | *pkd-2* | *ser-1* | *ser-4* | *tph-1* | *tub-1* | *tubby* | *unc-47* | *unc-49*
OMIM:
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM
Alzheimer's disease | Huntington's disease | Parkinson's disease

FURTHER INFORMATION
InParanoid: http://inparanoid.cgb.ki.se/index.html
OrthoDisease: http://orthodisease.cgb.ki.se
Access to this links box is available online.