

## ***Caenorhabditis elegans*: UNA HISTÒRIA FASCINANT**

DIANA DALFÓ

*Department of Pathology, Helen and Martin Kimmel Center for Stem Cells Biology,  
Skirball Institute of Biomolecular Medicine, New York University School of Medicine*

Adreça per a la correspondència: Diana Dalfó. Department of Pathology, Helen and Martin Kimmel Center for Stem Cells Biology, Skirball Institute of Biomolecular Medicine, New York University School of Medicine. First Avenue, 540, 4a, laboratori 7. 10016 Nova York (EUA). Adreça electrònica: [dalfod01@med.nyu.edu](mailto:dalfod01@med.nyu.edu).

### **RESUM**

El nematode *Caenorhabditis elegans* va ser escollit com a model genètic fa més de trenta anys. Des de llavors ha passat de ser un organisme que estava en la foscor a ser un dels organismes model més importants avui dia. *C. elegans* reuneix moltes característiques que en faciliten l'ús en estudis de genètica i de desenvolupament. Hi ha diversos factors que han influït en la seva gran popularitat, entre altres, el fet que fos el primer organisme multicel·lular que va tenir tot el seu genoma seqüenciat, i que es puguin expressar marcadors fluorescents *in vivo*. El gran avantatge de *C. elegans* com a sistema biològic és que permet estudiar qualsevol efecte en el context d'un organisme sencer. A més, el fet que sigui un organisme senzill però que al mateix temps molts processos biològics hi estiguin conservats al llarg de l'evolució fa d'aquest cuc un organisme ideal per estudiar malalties humanes. Aquesta revisió presenta els aspectes més importants d'aquest organisme, la gran variabilitat de recursos disponibles i, en definitiva, el gran potencial que té *C. elegans* per ajudar a entendre molts aspectes de la biologia humana.

**Paraules clau:** *Caenorhabditis elegans*, organisme model, GFP, RNAi, malaltia humana.

### ***Caenorhabditis elegans*: AN AMAZING STORY**

### **SUMMARY**

The nematode *Caenorhabditis elegans* was chosen as a model genetic organism more than 30 years ago. These years have taken this organism from obscurity to a position as one of the major model organisms. *C. elegans* has several distinct characteristics that make it suitable for genetic and developmental experiments. Several different aspects have

helped make it popular, among others, the fact that it was the first multicellular organism with its entire genome sequenced and that fluorescent markers can be readily observed *in vivo*. As an experimental system, *C. elegans* offers a unique opportunity to study any effect in a whole-organism context. The conservation of disease pathways between *C. elegans* and higher organisms, together with its simplicity, make of this organism an important model for human diseases. This review describes the key features of this organism, the different resources available and, in summary, the potential of *C. elegans* to help to understand the human biology.

**Key words:** *Caenorhabditis elegans*, model organism, GFP, RNAi, human disease.

Our field has prospered and has come of age; what was once a joke organism... has now become a major experimental system for the study of development and developmental genetics.

SYDNEY BRENNER (1987)

## INTRODUCCIÓ

El 1965, Sydney Brenner va escollir el nematode *Caenorhabditis elegans* (*C. elegans*) com a model per estudiar processos de desenvolupament i de neurobiologia. Avui, *C. elegans* s'ha convertit en «el cuc» i és amb tota probabilitat un dels organismes més ben conegut a escala anatòmica, genètica, de desenvolupament i de comportament. De fet, hi ha molts estudis que utilitzen *C. elegans* per investigar una gran varietat de processos com ara la regulació genètica, el metabolisme, el cycle celular, la mort celular programada, etc.

*C. elegans* és un cuc petit (1 mm de llarg), cilíndric, que viu a terra i que es pot trobar fàcilment en qualsevol part del món (vegeu la figura 1). S'alimenta de bacteris i el seu cycle vital —temps necessari per desenvolupar-se d'ou a adult— és aproximadament de tres dies. *C. elegans* és hermafrodita, és a dir, el mateix organisme produeix esperma i oòcits i es reproduïx per autofecundació. Tanmateix, alguns cucs mascle apareixen espontàniament, amb una freqüència molt baixa, i poden fertilitzar els cucs hermafrodites. Cada hermafrodita pot

produir fins a 300 ous al llarg de la seva vida reproductora. El desenvolupament d'aquest organisme consta de tres fases: embrió, larva i adult (Wood, 1988a) (vegeu la figura 2). La primera fase s'inicia amb la fecundació dins de l'úter d'un adult i es finalitza a l'exterior. A mig camí de completar el procés d'embriogènesi, l'embrió és esfèric i conté unes 500 cèl·lules. A partir d'aquest moment la proliferació celular s'atura i el cos s'allarga, el sistema nerviós es desenvolupa i finalment la primera fase larval (L1) trenca la cutícula, es mou lliurement i comença a alimentar-se de bacteris. Si les condicions ambientals no són bones (manca de menjar, temperatures altes, etc.) s'aturarà el desenvolupament i s'entrarà en una fase anomenada *dauer* (que en alemany significa 'permanent'). Com a *dauer* l'animal pot aguantar fins a quatre mesos en condicions molt adverses i sense la necessitat de menjar. Quan les condicions ambientals milloren, prosseguirà el desenvolupament larval, que consta de tres etapes més (L2, L3 i L4). Durant aquestes fases larvals l'organisme augmenta de mida, però el més important és que es desenvolupa la línia germinal que acabarà donant

lloc a l'esperma i als oòcits (Hubbard i Greenstein, 2005). Una cèl·lula específica de l'embrió serà determinada genèticament per donar lloc a tota la línia germinal i això la farà diferent de la resta de les cèl·lules, que es diferenciaran en teixits somàtics, com ara la hipodermis, l'intestí, el sistema nerviós, etc. Aquesta cèl·lula germinal tindrà un procés de proliferació molt important en la fase larval, i passarà d'una sola cèl·lula a gairebé 2.000 cèl·lules. Aquest creixement s'evidenciarà per dos braços que naixeran a cada costat de l'úter i que s'aniran allargant a mesura que les cèl·lules germinals proliferin fins que formin dos braços en forma de *u* (vegeu la figura 1). Fins al final de la L3 totes les cèl·lules que formen part dels dos braços de la gònada estan en cicle mitòtic i es caracteritzen per multiplicar-se constantment. A partir d'a-

quest moment, però, serà la primera vegada que les cèl·lules que estan més a prop de l'úter (a la part proximal de cada braç) entraran en meiosi, i, per tant, en altres paraules, es diferenciaran (ja no podran multiplicar-se més) i donaran lloc als gàmetes, primer esperma i després oòcits. En la fase L4 es formarà la vulva, que unirà l'úter amb l'exterior (Sulston, 1988). En l'última fase del cicle vital, en l'adult, ja no hi haurà més divisions de cèl·lules somàtiques, però les cèl·lules germinals en mitosi es continuaran dividint al llarg de la vida adulta (Hirsh *et al.*, 1976), la qual cosa les fa un model molt interessant per estudiar cèl·lules mare (Hubbard, 2007). Finalment, tindrà lloc la fertilització, cosa que tanca el cicle vital de *C. elegans*.

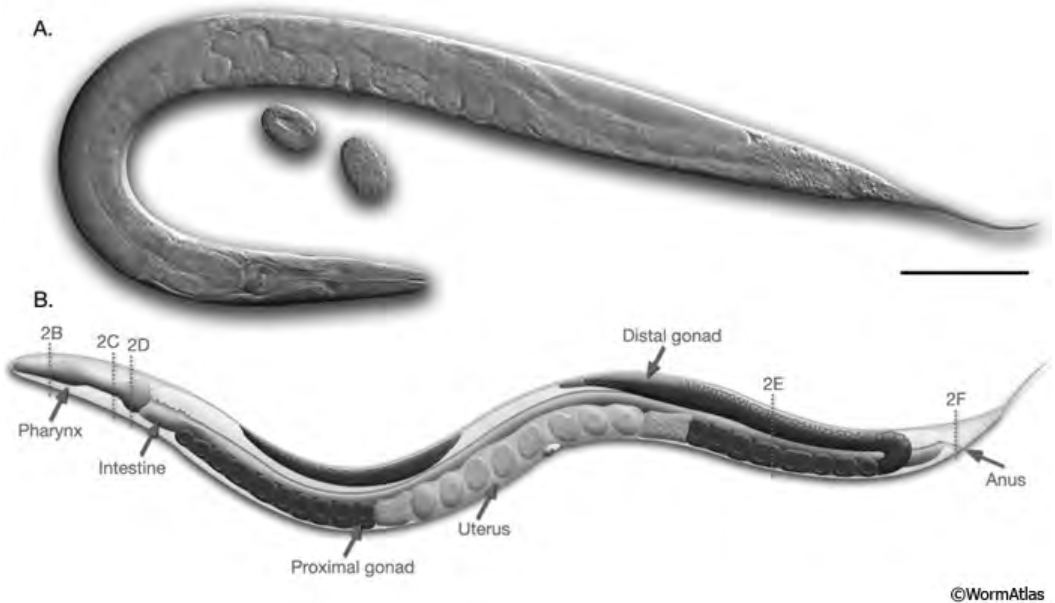


FIGURA 1. Anatomia d'un hermafrodita adult. A) Imatge DIC del costat esquerre d'un adult i de dos embrions que han estat expulsats a l'exterior (barra: 0,1 mm). B) Dibuix esquemàtic de les estructures anatòmiques d'un adult: faringe, intestí, úter, part proximal i distal de la gònada en forma de *u* i anus (©wormatlas.org).

### Quines característiques té *Caenorhabditis elegans* que el fan un bon organisme model?

a) És fàcil de mantenir amb una dieta d'*Escherichia coli* (*E. coli*), el bacteri més predominant en un laboratori. Només s'ha d'afegir una capa petita de bacteri en una placa de NGM (*nematode growth medium*) i transferir-hi els cucs. La temperatura ideal de manteniment és entre 16-25 °C. El ritme de creixement dependrà de la temperatura, i és 2,1 vegades més ràpid a 25 °C que a 16 °C. Aquesta diferència pot ser molt útil quan es plantegen diversos experiments.

b) *C. elegans* es reproduïx ràpidament. A més, el fet que en només tres dies es desenvolupi d'ou a adult, que puguem obtenir 300 ous per hermafrodita i que, com que les plaques ocupen poc espai, en puguem acumular tantes com vulguem, fa possible generar milions d'animals al dia.

c) Gràcies a la seva mida petita molts experiments es poden dur a terme en plaques de 96 pouets, cosa que permet analitzar un gran nombre d'animals per cada experiment i redueix significativament el cost de logística. A més a més, l'ús d'aquestes plaques ha fet possible la introducció de robots per manipular les mostres.

d) És transparent al llarg de tot el desenvolupament. Això implica que amb un microscopi podem observar divisions cel·lulars, migracions i processos de diferenciació en un animal viu. Si, a més, fem servir algun marcador fluorescent, es poden estudiar *in vivo* fenòmens com ara el creixement dels axons, el metabolisme dels greixos o l'embriogènesi mateixa.

e) És un animal multicel·lular sofisticat amb formació d'òrgans i teixits, com ara múscul, hipodermis, intestí, sistema reproductor i sistema nerviós.

Un dels motius més importants pel qual s'utilitzen organismes model és que molts

estudis en humans, especialment estudis genètics, són extremadament difícils de portar a terme a causa de la complexitat de la biologia humana, la falta d'eines per alterar l'activitat específica d'un gen *in vivo* i, evidentment, tot de qüestions ètiques que aquests experiments comportarien. *C. elegans*, en canvi, és un model ideal per a estudis genètics, ja que el fet que s'autofecundi permet obtenir fàcilment individus homozigots per a una mutació específica (revisat a Jorgensen i Mango, 2002). A més, el fet que tant en humans com en formes més simples es conservin vies metabòliques senceres i complexos de proteïnes similars, fa que molts laboratoris utilitzin aquests organismes per establir models funcionals que ens poden ajudar a entendre la biologia humana.

Sydney Brenner va concebre inicialment *C. elegans* com un sistema model per entendre aspectes fonamentals de genètica, desenvolupament i comportament. Brenner estava interessat en la relació entre gens i comportament però considerava que la mosca del vinagre, estudiada en gran detall a escala genètica, era massa complexa per a aquesta tasca. Brenner volia un organisme més senzill del qual es poguessin determinar totes les estructures del sistema nerviós. Al cap d'uns anys de la seva proposta es va aconseguir el que havia predit: *C. elegans* va ser el primer organisme del qual es va descriure la pauta de totes les connexions neuronals, que van ser reconstruïdes amb l'ajuda d'unes seccions en sèrie d'un microscopi electrònic (White *et al.*, 1986). A més, el fet que les divisions cel·lulars siguin invariables entre individus ha fet possible estudiar el llinatge de totes les cèl·lules somàtiques d'aquest organisme (959 cèl·lules). Partint del zigot, que és una única cèl·lula, s'ha pogut seguir el naixement de totes les cèl·lules, quantes vegades proliferen, si tenen la capacitat de migrar i

com es diferenciaran, és a dir, a quin teixit donaran lloc. Encara que l'anatomia d'aquest cuc és senzilla, se sap que les 959 cèl·lules somàtiques de l'adult representen la major part de teixits diferenciats que podem trobar en humans, com ara múscul (11 cèl·lules), neurones (302 cèl·lules), intestí (34 cèl·lules) i epidermis (213 cèl·lules) (Sulston i Horvitz, 1977). Finalment, s'ha de destacar que el primer organisme multicel·lular que va tenir el seu genoma completament seqüenciat va ser, naturalment, el cuc *C. elegans* (The *C. elegans* Consortium, 1998). Sydney Brenner va tenir una gran intuïció quan va escollir aquest petit animal com a organisme model, ja que ha

quedat palès que el seu gran potencial no es limita a aquests aspectes sinó que es pot utilitzar per estudiar una gran varietat de processos biològics, incloent-hi diverses malalties humanes. En reconeixement a la gran aportació que Sydney Brenner va fer a la comunitat científica va ser guardonat amb el Premi Nobel de Medicina l'any 2002. Actualment, més de 35 anys després del primer article publicat per Sydney Brenner, hi ha més de 700 laboratoris registrats que utilitzen *C. elegans* com a organisme model de manera rutinària.

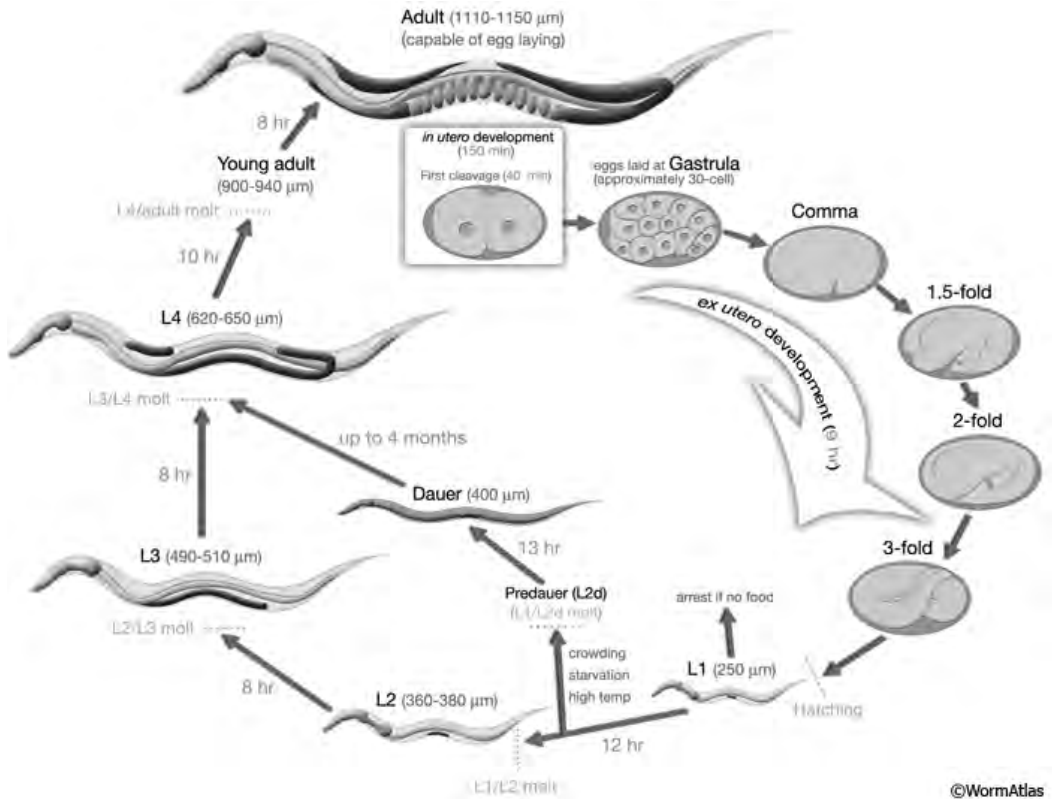


FIGURA 2. Cicle vital de *C. elegans* a 22 °C. El moment de la fecundació és considerat el minut zero. La durada de cada etapa està indicada en hores (hr). La llargada de l'animal està indicada en micròmetres (µm) (©wormatlas.org).

## TÈCNiques DISPONIBLES EN *Caenorhabditis elegans*

Una de les primeres coses que sorprenen més quan es comença a treballar amb aquest curiós cuc és la gran varietat de tècniques desenvolupades, com ara tècniques clàssiques de genètica, de biologia molecular, d'introducció de DNA o RNA exogen en la línia germinal del cuc, etc. S'ha de destacar, però, que dos dels avenços més importants han estat guardonats també amb el Premi Nobel, com a reconeixement al gran paper que han tingut no solament dins del camp dels nematodes sinó de la ciència en general. Un d'aquests descobriments va ser la introducció de marcadors fluorescents i l'altre es refereix a l'RNA d'interferència.

### *Green fluorescent protein*

Una de les històries més captivadores sobre aquest organisme fa referència al descobriment de la proteïna fluorescent verda o GFP (de les sigles en anglès, *green fluorescent protein*). El 1961 Osamu Shimomura es va posar a treballar amb la medusa *Aequorea victoria* amb la idea d'estudiar-ne la bioluminescència. L'objectiu era clar: purificar la substància luminescent que permetia a aquesta medusa emetre una fascinant llum blava. Tanmateix, tots els intents del jove Shimomura van fracassar. Va treballar dia i nit, va provar centenars de tècniques diferents, però cada vegada que intentava extreure aquella substància fluorescent acabava amb una gran foscor. Un dia, desesperat després d'un altre intent sense èxit, va llançar l'extracte que havia obtingut a la pica i llavors va quedar sorprès en veure que una intensa llum blava il·luminava la pica per uns segons. Aquella pica s'estava utilitzant com a desguàs d'un

aquari d'aigua salada situat prop del laboratori. Shimomura va deduir que el calci de l'aigua salada havia fet possible la luminescència. A partir d'aquí va ser més fàcil purificar aquesta nova proteïna, que en presència de calci emetia llum blava. La va anomenar *aequorina* (Shimomura *et al.*, 1962). Tanmateix, a més d'haver-hi aequorina en la medusa també hi havia present una proteïna més petita que en algunes circumstàncies podia rebre la llum blava produïda per l'aequorina i emetre llum verda. Quan Shimomura va purificar l'aequorina també va purificar aquesta proteïna, i com que emetia llum verda, la va anomenar GFP (Morise *et al.*, 1974). La GFP, però, tenia una característica que la diferenciava de qualsevol altra proteïna fluorescent. Fins aquell moment es creia que les proteïnes fluorescents eren el resultat d'un complex format per una proteïna i una substància fluorescent. La GFP era una proteïna fluorescent molt especial perquè no necessitava cap substància fluorescent per emetre llum, com l'aequorina, ja que una part intrínseca de la seva molècula és la responsable d'emetre llum. Aquest descobriment va ser molt important, ja que suggeria que

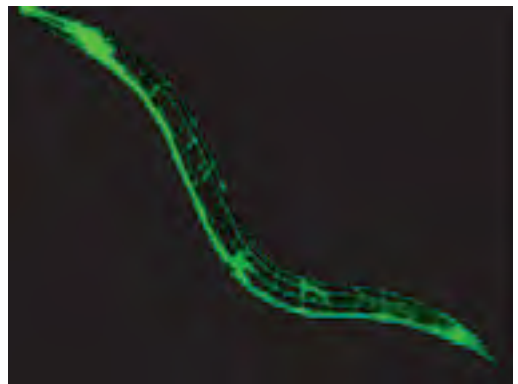


FIGURA 3. Expressió de la GFP en *C. elegans* *in vivo*. Adult hermafrodita que expressa GFP en totes les cèl·lules del sistema nerviós (Chalfie, 2009).

es podia clonar la GFP i que hauria de ser capaç d'emetre llum per si sola. Tanmateix, Shimomura va deixar de treballar amb la GFP i no es va saber res més d'aquesta proteïna fins al 1989. Martin Chalfie, professor de la Columbia University, a Nova York, va assistir a una conferència en què es parlava dels descobriments fets més de vint anys enrere per Shimomura. En aquell temps Chalfie estava treballant en l'expressió de diversos gens en el sistema nerviós del *C. elegans*. En aquella època hi havia diferents sistemes per detectar l'expressió de gens i proteïnes, però tots obligaven a fixar l'animal, i per tant només es podien observar teixits d'animals morts. Quan Chalfie va sentir per primera vegada parlar d'una proteïna que emetia llum verda de seguida va començar a imaginar de quina manera la podia fer servir com a marcador biològic per estudiar cucs vius, els quals, recordem-ho, són transparents al llarg de tot el seu cicle vital (vegeu la figura 3). El 1994 Chalfie i els seus col·laboradors van publicar un article a la revista *Science* en què demostraven que l'expressió de la GFP en organismes procariotes (*E. coli*) o eucariotes (*C. elegans*) produïa una intensa llum verda que es podia fer servir per seguir l'expressió de gens o la localització de proteïnes en organismes vius (Chalfie *et al.*, 1994). La importància que ha tingut la GFP en l'estudi de processos *in vivo* ha estat reconeguda amb el Premi Nobel de Química de l'any 2008, quan Osamu Shimomura juntament amb Martin Chalfie i Roger Tsien van ser guardonats amb aquest prestigiós premi. Avui, molts laboratoris utilitzen la GFP de manera rutinària amb diverses finalitats; per exemple, com a marcador de proteïnes, de compartiments subcel·lulars, de cèl·lules, per seguir processos de migració o per seguir la introducció de DNA extern en *C. elegans* (revisat a Hodgkin i Herman, 1998). Actualment hi ha diverses soques comerci-

als que expressen la GFP en diversos teixits o tipus cel·lulars.

## RNA d'interferència

Cap a la meitat dels anys cinquanta es va introduir per primera vegada el concepte de RNA de doble cadena (dsRNA, de les sigles en anglès *double-strand RNA*): quan dues cadenes de RNA mostren grans regions de seqüència complementària poden unir-se (com si fossin les dues tires d'una cremallera) formant una molècula de dsRNA (Rich i Davies, 1956). Al principi dels anys seixanta, però, es creia que la font d'aquest tipus d'estructura era majoritàriament vírica (Baltimore *et al.*, 1964). En aquella època el dogma deia que el DNA de doble hèlix, on es guarda la informació genòmica, es transcriu a RNA d'una sola cadena, que al seu torn es tradueix a proteïna. En aquell moment el dsRNA no tenia cap paper en la biologia cel·lular. No va ser fins al final dels anys noranta quan Craig Mello i Andrew Fire van entrar en escena. El seu objectiu era disminuir l'expressió d'un gen utilitzant RNA d'una sola cadena. L'RNA era introduït en la gònada del cuc amb una agulla molt petita, tècnica que rep el nom de *microinjecció* (Mello *et al.*, 1991). Aquests investigadors van començar utilitzant l'RNA complementari al gen *unc-22*, un dels responsables de controlar el moviment del cuc. La idea era que quan l'RNA introduït es trobés una molècula de RNA endògena del gen *unc-22*, com que les seqüències serien complementàries, s'unirien formant una cremallera i això n'impediria la traducció a proteïna i disminuiria així l'activitat del gen en qüestió. L'efecte, tot i que evident, no era gaire significatiu, ja que només una petita part de la descendència mostrava un moviment deficient. En aquell moment ja se sabia que el

dsRNA era bastant més estable que l'RNA d'una sola cadena i en un moment d'inspiració aquests investigadors van decidir barrejar en un tub una cadena de RNA sentit i una d'antisentit per generar així molècules de dsRNA. Van intentar de nou les microinjeccions, però aquesta vegada amb dsRNA. Per a sorpresa seva l'efecte va ser molt més gran i ara la major part de la descendència tenia un moviment defectuós (Fire *et al.*, 1998). Havia nascut l'RNA d'interferència (RNAi). S'havia demostrat que el dsRNA era capaç de degradar l'RNA missatger (mRNA) de la cèl·lula amb la qual comparteix una seqüència idèntica. Més tard es va diversificar la manera d'introduir el dsRNA. Tot i que la microinjecció funcionava, era una tècnica que requeria una gran dosi d'experiència i molta habilitat. Així doncs, es va decidir modificar genèticament el bacteri que es menja el cuc, l'*E. coli*, perquè pogués generar dsRNA a partir d'un vector que conté una seqüència de DNA. Per a sorpresa seva, quan el *C. elegans* es menjava aquest bacteri es veia el mateix efecte en la descendència que amb les microinjeccions (Timmons i Fire, 1998). També van demostrar que si se submergia un cuc en una solució que contingués molècules de dsRNA aquestes eren capaces de travessar les membranes cel·lulars i distribuir-se per tot el cos (Tabara *et al.*, 1998). D'una manera que avui encara no està del tot clara, el dsRNA és capaç de distribuir-se al llarg del cos del cuc, independentment de com hagi entrat. Tots aquests mecanismes produeixen una disminució de l'expressió d'un gen específic que es pot trobar en qualsevol lloc de l'organisme. La gran aportació d'aquests dos científics va ser l'estudi dels gens implicats en el mètode pel qual el dsRNA és capaç de silenciar l'expressió d'un gen (Parrish *et al.*, 2000). El fet que es trobessin resultats similars en mosques, plantes i finalment

humans va fer evident que es tractava d'un procés conservat evolutivament. És a dir, el mecanisme amb el qual es basa l'RNAi és, de fet, molt antic, està present de manera natural en diversos organismes i s'utilitza com a defensa davant de virus o transposons. Avui, l'RNAi s'ha convertit en una de les tècniques més utilitzades per estudiar la funció d'un gen, especialment en *C. elegans* (revisat a Boutros i Ahringer, 2008). Els avantatges de l'RNAi són diversos, com ara el fet que la seqüència dels gens implicats és coneguda abans de començar l'experiment; el dsRNA es pot introduir en diferents estadis del desenvolupament i es poden silenciar famílies gèniques senceres si comparteixen un fragment de seqüència, i s'elimina així el problema de la redundància funcional. D'altra banda, no tots els gens són receptius a l'RNAi, alguns teixits no responen a aquesta tècnica (especialment el teixit neural), l'efecte que produeix l'RNAi desapareix quan s'elimina la font del dsRNA (per exemple, menjar) i finalment s'ha de tenir en compte que de vegades no s'aconsegueix una reducció total del nivell de l'mRNA, i per tant encara hi pot haver una petita producció de proteïna. Com a reconeixement a la gran feina que van fer, Craig Mello i Andrew Fire van rebre el Premi Nobel el 2006.

## TÈCNiques I RECURSOS DISPONIBLES EN *Caenorhabditis elegans*

Els avenços que han tingut lloc en el camp del *C. elegans* han estat, en part, gràcies al fet que històricament hi ha hagut una gran col·laboració entre els diferents laboratoris. Això ha significat que hi hagi un gran nombre de tècniques i de recursos informàtics a l'abast de qualsevol investigador. A continuació hi ha un resum de les



pàgines d'Internet de més utilitat (revisat a Antoshechkin i Sternberg, 2007):

Els llibres *The nematode Caenorhabditis elegans* i el seu successor *C. elegans II* són d'un interès enorme (Riddle *et al.*, 1997; Wood, 1988b). Recentment es va decidir actualitzar una vegada més el llibre, però aquest cop amb una versió en línia (<http://www.wormbook.org>). L'objectiu és presentar de manera intel·ligible tots els aspectes coneguts de la biologia del cuc i descriure'n les diverses tècniques disponibles.

Cal destacar la formació recent d'una pàgina molt útil per a estudiants i professors de secundària (<http://www.wormclassroom.org>). En aquesta pàgina s'explica què és un organisme model, es dona una descripció general del *C. elegans* i a més hi ha accés a vídeos, protocols experimentals i molts recursos educatius més.

La base de dades <http://www.wormbase.org> és coneguda com la més important sobre *C. elegans* i altres nematodes. Conté informació sobre seqüències genòmiques i proposa models estructurals de gens. Cada gen té una pàgina pròpia on es resumeix tota la informació coneguda que s'extreu manualment dels articles publicats cada any. A més, el fet que el genoma d'altres nematodes estigui disponibles fa d'aquesta pàgina una eina molt important per a la genòmica comparada.

A <http://www.wormatlas.org> es dona una informació exhaustiva de l'anatomia del *C. elegans*. Cada capítol està acompanyat d'uns diagrames que descriuen les diferents parts anatòmiques amb gran detall. És una manera excel·lent d'explorar aquest organisme.

El *Caenorhabditis* Genetics Center (<http://biosci.umn.edu/CGC>) és el responsable de col·leccionar, mantenir i distribuir totes les soques que s'han generat de *C. elegans* des del 1978. El fet que es puguin congelar cucs en estadi larval (L1) permet tenir una

font permanent de totes les mutacions que s'han generat al llarg del temps. Una altra funció essencial del CGC és la coordinació de la nomenclatura. Històricament els investigadors en *C. elegans* han seguit normes molt estrictes per a cada gen descobert. Això ha assegurat una uniformitat dels noms dels gens que ha disminuït substancialment les confusions que poden donar-se en altres organismes quan el mateix gen rep dos noms diferents.

Textpresso (<http://textpresso.org>) és un cercador que va ser específicament dissenyat per a la bibliografia del *C. elegans*. Acumula més de vuit mil articles, a més de resums de xerrades dels congressos internacionals i nacionals. Cada any s'hi afegeixen al voltant d'uns mil articles. El gran avantatge d'aquest cercador és que pot accedir a tot el text d'aquests articles i xerrades, mentre que els altres cercadors es limiten al resum que hi ha davant de cada article.

Històricament el servidor *C. elegans* WWW (<http://elegans.swmed.edu>) ha servit per unir els diversos tipus d'informació que hi ha sobre *C. elegans* i com a punt de trobada entre els científics. A més a més, totes aquestes pàgines tenen enllaços entre si, cosa que en facilita la interconnexió.

## UN GENOMA COMPLETAMENT SEQÜENCIAT

Un dels grans avenços en la història d'aquest cuc va tenir lloc al final de l'any 1998, quan el *C. elegans* es va convertir en el primer organisme multicel·lular que va tenir tot el genoma seqüenciat (The *C. elegans* Consortium, 1998). Per poder completar el projecte es van necessitar uns 3.000 clons i la feina coordinada de centenars de científics del Sanger Institute (Regne Unit) i del Genome Sequencing Center (EUA). Aquest

èxit va ser un pas important per assolir la seqüència del genoma humà uns anys més tard (International Human Genome Consortium, 2004). Aquests dos fets van fer una empenta a la genòmica comparada, que es dedica a comparar els genomes de diverses espècies per estudiar l'efecte que tenen els processos evolutius. Avui sabem aproximadament quants gens hi ha en els humans i en organismes model i com estan distribuïts al llarg del genoma. El repte següent, però, serà entendre la funció de tots els gens i de les parts del genoma no codificants. En altres paraules, si ens imaginem que el genoma és un llibre i que cada base del DNA és una lletra, actualment coneixem totes les lletres que formen aquest llibre, la seva posició i que formen paraules. Sabem el significat d'algunes paraules però encara n'hi ha moltes que no sabem què volen dir. Si comparem diversos llibres podrem obtenir informació que ens ajudarà a entendre el significat de més paraules i de l'estructura dels llibres en general.

El genoma del *C. elegans* conté més de 19.000 gens que codifiquen proteïnes i que es troben dispersos al llarg dels 97 milions de bases de què consta el seu genoma. Aproximadament el 42 % de les proteïnes en *C. elegans* tenen alguna semblança amb proteïnes d'espècies que no són nematodes, la qual cosa suggereix que s'han conservat al llarg de l'evolució. Per contra, un 34 % de les proteïnes només mostren algun grau de semblança amb proteïnes que es troben exclusivament en els nematodes (Green *et al.*, 1993). Els científics s'han centrat en el primer grup, especialment si mostren homologia amb proteïnes humanes. El genoma, però, no està format exclusivament per gens que codifiquen proteïnes sinó que hi ha altres tipus de seqüències. En el cas del *C. elegans* s'han trobat diversos centenars de gens que es transcriuen a RNA però que no codifiquen

proteïnes, com per exemple els gens que codifiquen tRNA. A més, també hi ha seqüències no codificants que són importants per a la regulació gènica i per al manteniment dels cromosomes. Com en altres organismes multicel·lulars, una part important d'aquestes seqüències és repetitiva. Les repeticions en tàndem constitueixen el 2,7 % del genoma i les invertides un 3,6 %. No estan distribuïdes uniformement, ja que és més probable que es trobin en els introns que entre els gens (The *C. elegans* Consortium, 1998). S'ha de destacar que el genoma del cuc sembla bastant uniforme pel que fa al contingut de GC, que és pràcticament invariable al llarg de tots els cromosomes (al voltant d'un 36 %), a diferència del contingut de GC en vertebrats (Bernardi, 1995). A més, té un percentatge de seqüència codificant més alt que en l'humà, els introns i les seqüències reguladores són més petites i conté menys heterocromatina (més definició) que el genoma humà (Reinke i White, 2002). Una dada molt interessant que es va obtenir a partir de l'estudi del genoma del *C. elegans* és que aquest organisme conté un nombre de gens més alt del que s'havia predit. Com a referència tenim el genoma humà que, tot i contenir més de tres mil milions de bases de DNA (trenta vegades més gran que el genoma del cuc), «només» conté uns 30.000 gens (1,5 vegades més que el nombre de gens en *C. elegans*) (International Human Genome Consortium, 2004). Per què un animal suposadament «senzill» necessita uns 20.000 gens, si per fer un ésser humà sembla que amb 30.000 ja n'hi ha prou? Hi ha diverses raons que podrien explicar, almenys en part, per què aquestes dues espècies tenen un nombre de gens tan similar (revisat a Hodgkin, 2001). S'ha de tenir en compte que en el genoma humà hi ha una gran varietat de proteïnes gràcies a l'empalmament alternatiu (de l'anglès *al-*

ternative splicing). L'empalmament consisteix a eliminar les seqüències intròniques del pre-mRNA i unir els exons que estan separats pels introns per tal de formar un mRNA madur que serà traduït a proteïna. L'empalmament alternatiu és una sofisticació del sistema en el qual alguns exons també poden ser eliminats juntament amb els introns fronterers. Això incrementa significativament les diferents formes de mRNA, i per tant, de proteïnes, que es poden obtenir d'un sol gen. Dels 30.000 gens que hi ha en els humans, de fet, es poden arribar a generar al voltant d'unes 100.000 proteïnes diferents. En *C. elegans*, tot i que aquest fenomen també s'hi dona, és en una freqüència molt més baixa que en humans. Només s'ha pogut confirmar que succeeixi en un 4 % dels gens. Una altra diferència és que els gens humans tendeixen a codificar proteïnes que tenen diversos dominis funcionals, mentre que les proteïnes en *C. elegans* són generalment més senzilles i això fa necessari que diferents dominis siguin codificats per gens independents. Un altre factor important és que hi ha hagut diversos processos de duplicacions específics de la línia dels nematodes, i això fa que molts gens en el cuc s'hagin duplicat. Això genera famílies gèniques grans que poden dificultar els estudis genètics, ja que no seria suficient mutar un sol gen per anul·lar la funció i s'hauria de recórrer a mutants dobles o més complexos. S'ha de dir, però, que en el cas del *C. elegans*, en la majoria dels casos la redundància funcional és incompleta. És a dir, dos gens poden coincidir en algunes de les seves funcions, però cada un té com a mínim una funció específica que no té l'altre. Un bon exemple d'això el trobem en els dos receptors de Notch en el cuc, *lin-12* i *glp-1*, que comparteixen algunes funcions durant l'embriogènesi però tenen funcions postembriòniques diferents (Kimble i Simpson, 1997).

El 2002 es va poder accedir per primera vegada al genoma complet d'un altre nematode, el *C. briggsae* (Stein *et al.*, 2003). Estudiar les relacions filogenètiques entre *C. elegans* i altres nematodes és important per a l'anàlisi comparada de comportament, morfologia, desenvolupament, mecanismes moleculars i de genomes (Gupta, 2007). Tot i que aquestes dues espècies van divergir d'un ancestre comú fa més de 100 milions d'anys, tenen un desenvolupament i una morfologia molt similar. De fet, comparteixen un 80 % d'identitat en seqüències que estan sota una gran pressió per conservar-se però només un 30 % d'identitat en altres parts del genoma (Kent i Zahler, 2000). L'estudi comparatiu entre *C. briggsae* i *C. elegans* pot permetre inferir característiques específiques de cada espècie, però també pot donar informació sobre quines característiques s'han conservat al llarg de l'evolució i quines s'han quedat restringides als nematodes.

### *C. elegans* i malalties humanes

Es pot utilitzar *C. elegans* com a organisme model per estudiar malalties humanes? El millor organisme per estudiar com es desenvolupen malalties i per analitzar tractaments és, segurament, el ratolí, ja que representa el model més semblant que tenim a l'ésser humà. Tanmateix, l'estudi d'organismes més senzills pot ser d'una gran ajuda per estudiar funcions bàsiques dels gens o per identificar els membres implicats en una via metabòlica. Una de les raons utilitzades en contra de l'ús del *C. elegans* en aquest camp és la idea que un cuc i un humà són tan diferents que no és possible que tinguin res en comú. Gràcies a la genòmica comparada s'ha identificat que entre un 60-80 % dels gens implicats en alguna malaltia humana tenen un gen homò-

leg en *C. elegans* (Lai et al., 2000; Culetto i Sattelle, 2000). Aquests estudis van revelar una gran i sorprenent conservació a escala molecular i de processos cel·lulars entre l'ésser humà i el cuc que permet l'ús del *C. elegans* en aquests estudis. A més, s'ha de recordar que, malgrat que el cuc és un organisme «senzill», té una fisiologia basada en la formació d'òrgans, com la que té l'ésser humà. Per tant, *C. elegans* pot ser un model per a malalties molt valuós sempre que la malaltia es pugui definir a escala molecular. La genòmica comparada també ha servit per identificar en *C. elegans* els ortòlegs de molts gens implicats en malalties humanes (els ortòlegs són gens que es troben en dues espècies i que han evolucionat directament a partir d'un sol gen que es trobava en l'ancestre comú). D'acord amb el grau de conservació de seqüència entre dos gens es pot predir la probabilitat que comparteixin la mateixa funció. Malauradament, l'evolució d'aquesta no sempre és paral·lela a l'evolució de la seqüència. Per tant, per decidir si *C. elegans* és un bon model per a una malaltia específica no s'hauria d'utilitzar el grau de conservació de la seqüència com a criteri únic. Com a exemple tenim el receptor GABA, que està implicat en l'epilèpsia i que, per tant, és un destinatari important de les drogues antiepilèptiques. Aquesta funció està conservada independentment del grau de conservació de la seqüència. En un primer moment els programes informàtics per identificar ortòlegs en *C. elegans* no van identificar el gen *unc-49* com a ortòleg del receptor humà GABA, la qual cosa hauria suggerit que *C. elegans* no era un bon model per estudiar l'epilèpsia. Tanmateix, estudis d'aquest gen en el cuc han aportat molta informació sobre la ruta metabòlica GABA. És cert que la situació contrària també existeix: en una família gènica gran, com per exemple la família dels receptors

*nhr* (de l'anglès *nuclear hormone receptors*), és molt difícil d'assenyalar quin gen és l'autèntic ortòleg, encara que la seqüència entre el cuc i l'ésser humà estigui altament conservada (revisat a Kaletta i Hengartner, 2006). Una altra de les limitacions en l'ús d'aquest organisme és que no es pot esperar que pugui recapitular tots els símptomes característics de les malalties humanes, però els estudis de genètica, tan fàcils en el cuc i gairebé tan impossibles en l'ésser humà, ajudaran a tenir més informació sobre la base genètica de les malalties humanes. A continuació es donen alguns exemples que demostren que aquest organisme pot ser molt útil per a aquest tipus d'estudis si se'n coneixen els avantatges i les limitacions com a organisme model.

El gen humà *DPC4*, implicat en el càncer del pàncrees, és similar al gen *sma-4* de *C. elegans*. Els experiments fets amb aquest organisme demostren que aquest gen forma part de la ruta metabòlica del TGF- $\beta$  (*transforming growth factor  $\beta$* ), que controla el creixement (Savage et al., 1996). Per tant, una mutació en el gen *DPC4* podria causar càncer per una alteració en el senyal del TGF- $\beta$ . Més estudis en *C. elegans* poden donar informació directa de com funcionen els gens humans.

*C. elegans* també ha tingut un paper important en els estudis sobre la malaltia d'Alzheimer. La proteïna humana pèptid  $\beta$ -amiloide, que està implicada en aquesta malaltia, ha estat expressada amb èxit en *C. elegans* (Link, 1995). Prèviament als estudis fets amb *C. elegans* es considerava que l'acumulació del pèptid  $\beta$ -amiloide era un marcador per a la malaltia i segurament era el primer pas per desenvolupar-la. Diversos estudis fets amb el cuc van demostrar que aquesta acumulació és més aviat el resultat d'una resposta cel·lular per eliminar proteïnes que no són correctes i, per tant, és la conseqüència d'una disfunció

cellular, més que ser l'iniciador del problema (Cohen *et al.*, 2006).

L'obesitat és un dels problemes més greus del nostre temps i els pronòstics per al futur no són gaire bons. Tots els organismes han d'arribar a un equilibri entre l'energia que entra amb el menjar i la que s'utilitza per sobreviure. Moltes de les proteïnes implicades en la formació, el transport i l'ús dels greixos estan altament conservades entre els nematodes i els mamífers. L'ús del *C. elegans* ha permès la identificació de nous gens necessaris per a l'equilibri de l'energia. A més, el fet que el sistema nerviós d'aquest organisme sigui més senzill que l'humà el fa un sistema atractiu per estudiar com els organismes detecten nutrients i canvien el seu comportament per adaptar-s'hi. Depenent de si hi ha menjar o no, els cucs alteraran la velocitat amb la qual s'alimenten, amb la qual es mouen, i també el nombre d'ous generats. Molts dels components necessaris per a aquests canvis de comportament també són responsables del comportament de l'ésser humà respecte al menjar, com per exemple la serotonina (Jones i Ashrafi, 2009).

També s'ha utilitzat *C. elegans* per estudiar gens implicats en l'esperança de vida. S'han identificat uns cent gens que quan estan mutats incrementen l'esperança de vida del cuc. Entre aquests destaca el paper de la insulina (Hamilton *et al.*, 2009).

Tots els animals estan exposats a una gran varietat de microbis i han desenvolupat complexes respostes immunitàries per protegir-se. En els vertebrats hi ha dos tipus de respostes: la innata i l'adaptada. Mentre que la segona és específica dels vertebrats, la primera és la més universal i la que reacciona abans. *C. elegans* és un bon model per estudiar quins gens són importants per protegir-se dels patògens. A part de les barreres físiques com la hipodermis

o la faringe (extremadament eficaç en triturar microorganismes) hi ha com a mínim sis vies metabòliques que s'activen depenent del tipus d'infecció: TGF- $\beta$ , insulina, mort celular programada i tres vies diferents de MAP-cinases. També té un paper important el sistema nerviós, que permet al cuc detectar ambients no saludables i apartar-se'n per moure's cap a ambients més bons. Algunes d'aquestes respostes poden ser específiques del *C. elegans* però d'altres poden representar respostes conservades al llarg de l'evolució que ens poden ajudar a combatre infeccions en l'ésser humà (Gravato-Nobre i Hodgkin, 2005).

*C. elegans* també ha contribuït amb estudis fets sobre el Parkinson, la distròfia muscular, la malaltia de Huntington, el càncer, la diabetis, la depressió, etc. (revisat a Kaletta i Hengartner, 2006). Sobretot en aquest últim cas ha estat molt útil el desenvolupament dels cribratges amb components químics. Això ha permès analitzar el paper d'algunes drogues (com alguns antidepressius) i buscar mutacions de manera eficaç i ràpida que facin que el cuc sigui insensible o hipersensible a la droga. El gran avantatge d'utilitzar *C. elegans* és que permet estudiar *in vivo* els efectes del component químic en un organisme sencer i és suficientment simple perquè se'n pugui fer la dissecció del mecanisme d'acció amb tot detall. L'altre avantatge és que permet eliminar components tòxics en els primers estudis d'acció, i s'estalvien així els estudis en organismes més complexos (revisat a Ahringer, 1997).

Voldria acabar amb una reflexió personal. L'objectiu d'aquest capítol era mostrar el gran potencial del *C. elegans* com a organisme model i ensenyar les múltiples possibilitats que hi ha d'utilitzar-lo. Espero haver pogut transmetre com és de fascinant la història d'aquest «senzill» cuc i com és d'apassionant la ciència en general. Perso-

nalment crec que el fet que hi hagi tants processos conservats entre l'ésser humà i el cuc, i que tants aspectes de la vida es puguin estudiar en *C. elegans*, tot i que no tots, no deixa de ser una lliçó moral per a nosaltres, els humans. Finalment, espero que el nostre amic *C. elegans* s'hagi guanyat el respecte dels lectors.

## BIBLIOGRAFIA

- AHRINGER, J. (1997). «Turn to the worm!» *Curr. Opin. Genet. Dev.*, 7: 410-415.
- ANTOSHECHKIN, I.; STERNBERG, P. W. (2007). «The versatile worm: genetic and genomic resources for *Caenorhabditis elegans* research». *Nat. Rev. Genet.*, 8: 518-532.
- BALTIMORE, D.; BECKER, Y.; DARNELL, J. E. (1964). «Virus-specific double-stranded RNA in poliovirus-infected cells». *Science*, 143: 1034-1036.
- BERNARDI, G. (1995). «The human genome: organization and evolutionary history». *Annu. Rev. Genet.*, 29: 445-476.
- BOUTROS, M.; AHRINGER, J. (2008). «The art and design of genetic screens: RNA interference». *Nat. Rev. Genet.*, 9: 554-566.
- CHALFIE, M. (2009). «GFP: lighting up life». *Angew. Chem. Int. Ed. Engl.*, 48: 5603-5611.
- CHALFIE, M.; TU, Y.; EUSKIRCHEN, G.; WARD, W.; PRA-SHER, D. C. (1994). «Green fluorescent protein as a marker for gene expression». *Science*, 263: 802-805.
- COHEN, E.; BIESCHKE, J.; PERCIAVALLE, R. M.; KELLY, J. W.; DILLIN, A. (2006). «Opposing activities protect against age-onset proteotoxicity». *Science*, 313: 1604-1610.
- CULETTO, E.; SATTELLE, D. B. (2000). «A role for *C. elegans* in understanding the function and interactions of human disease genes». *Hum. Mol. Genet.*, 9: 869-877.
- FIRE, A.; XU, S.; MONTGOMERY, M. K.; KOSTAS, S. A.; DRIVER, S. E.; MELLO, C. C. (1998). «Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*». *Nature*, 391: 806-811.
- GRAVATO-NOBRE, M. J.; HODGKIN, J. (2005). «*Caenorhabditis elegans* as a model for innate immunity to pathogens». *Cell Microbiol.*, 7: 741-751.
- GREEN, P.; LIPMAN, D.; HILLIER, L.; WATERSTON, R.; STATES, D.; CLAVERIE, J. M. (1993). «Ancient conserved regions in new gene sequences and the protein databases». *Science*, 259: 1711-1716.
- GUPTA, B. P. (2007). «Genomics and biology of the nematode *C. briggsae*». *WORMBOOK* [ed.]. *The C. elegans Research Community, WormBook*. DOI: 10.1895/wormbook.1.136.1.
- HAMILTON, B. [et al.] (2005). «A systematic RNAi screen for longevity genes in *C. elegans*». *Genes Dev.*, 19: 1544-1555.
- HIRSH, D.; OPPENHEIM, D.; KLASS, M. (1976). «Development of the reproductive system of *Caenorhabditis elegans*». *Dev. Biol.*, 49: 200-219.
- HODGKIN, J. (2001). «What does a worm want with 20,000 genes?» *Genome Biol.*, 2: 1-4.
- HODGKIN, J.; HERMAN, R. K. (1998). «Changing styles in *C. elegans* genetics». *Trends Genet.*, 14: 352-357.
- HUBBARD, E. J. A. (2007). «*Caenorhabditis elegans* germ line: a model for stem cell biology». *Dev. Dyn.*, 236: 3343-3357.
- HUBBARD, E. J. A.; GREENSTEIN, D. (2005). «Introduction to the germ line». *WORMBOOK* [ed.]. *The C. elegans Research Community, WormBook*. DOI: 10.1895/wormbook.1.18.1.
- INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM. (2004). «Finishing the euchromatic sequence of the human genome». *Nature*, 431: 931-945.
- JONES, K. T.; ASHRAFI, K. (2009). «*Caenorhabditis elegans* as an emerging model for studying the basic biology of obesity». *Dis. Model Mech.*, 2: 224-229.
- JORGENSEN, E. M.; MANGO, S. E. (2002). «The art and design of genetic screens: *Caenorhabditis elegans*». *Nat. Rev. Genet.*, 3: 356-369.
- KALETTA, T.; HENGARTNER, M. O. (2006). «Finding function in novel targets: *C. elegans* as a model organism». *Nat. Rev. Drug Discov.*, 5: 387-398.
- KENT, W. J.; ZAHLER, A. M. (2000). «Conservation, regulation, synten, and introns in a large-scale *C. briggsae-C. elegans* genomic alignment». *Genome Res.*, 10: 1115-1125.
- KIMBLE, J.; SIMPSON, P. (1997). «The LIN-12/Notch signaling pathway and its regulation». *Annu. Rev. Cell. Dev. Biol.*, 13: 333-361.
- LAI, C. H.; CHOU, C. Y.; CHANG, L. Y.; LIU, C. S.; LIN, W. (2000). «Identification of novel human genes evolutionarily conserved in *Caenorhabditis elegans* by comparative proteomics». *Genome Res.*, 10: 703-713.
- LINK, C. D. (1995). «Expression of human  $\beta$ -amyloid peptide in transgenic *Caenorhabditis elegans*». *Proc. Natl. Acad. Sci. USA*, 92: 9368-9372.
- MELLO, C. C.; KRAMER, J. M.; STINCHCOMB, D.; AMBROS, V. (1991). «Efficient gene transfer in *C. ele-*

- gans*: extrachromosomal maintenance and integration of transforming sequences». *EMBO J.*, 10: 3959-3970.
- MORISE, H.; SHIMOMURA, O.; JOHNSON, F. H.; WINANT, J. (1974). «Intermolecular energy transfer in the bioluminescent system of *Aequorea*». *Biochemistry*, 13: 2656-2662.
- PARRISH, S.; FLEENOR, J.; XU, S.; MELLO, C.; FIRE, A. (2000). «Functional anatomy of a dsRNA trigger: differential requirement for the two trigger strands in RNA interference». *Mol. Cell.*, 6: 1077-1087.
- REINKE, V.; WHITE, K. P. (2002). «Developmental genomic approaches in model organisms». *Annu. Rev. Genomics Hum. Genet.*, 3: 153-178.
- RICH, A.; DAVIES, D. R. (1956). «A new two-stranded helical structure: polyadenylic acid and polyuridylic acid». *J. Am. Chem. Soc.*, 78: 3548-3549.
- RIDDLE, D. L.; BLUMENTHAL, T.; MEYER, B. J.; PRIESS, J. R. (1997). *C. elegans II*. Plainview: Cold Spring Harbor Laboratory Press.
- SAVAGE, C.; DAS, P.; FINELLI, A. L.; TOWNSEND, S. R.; SUN, C. Y.; BAIRD, S. E.; PADGETT, R. W. (1996). «*C. elegans* genes *sma-2*, *sma-3*, and *sma-4* define a conserved family of TGF- $\beta$  pathway components». *Proc. Natl. Acad. Sci. USA*, 93: 790-794.
- SHIMOMURA, O.; JOHNSON, F. H.; SAIGA, Y. (1962). «Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, *Aequorea*». *J. Cell. Comp. Physiol.*, 59: 223-239.
- STEIN, L. D. [et al.] (2003). «The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics». *PLoS. Biol.*, 1: e45.
- SULSTON, J. (1988). «Cell lineage». A: WOOD, W. B. [ed.]. *The Nematode Caenorhabditis elegans*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, p. 123-155.
- SULSTON, J. E.; HORVITZ, H. R. (1977). «Post-embryonic cell lineages of the nematode *Caenorhabditis elegans*». *Dev. Biol.*, 56: 110-156.
- TABARA, H.; GRISHOK, A.; MELLO, C. C. (1998). «RNAi in *C. elegans*: soaking in the genome sequence». *Science*, 282: 430-431.
- THE *C. elegans* SEQUENCING CONSORTIUM. (1998). «Genome sequence of the nematode *C. elegans*: a platform for investigating biology». *Science*, 282: 2012-2018.
- TIMMONS, L.; FIRE, A. (1998). «Specific interference by ingested dsRNA». *Nature*, 395: 854.
- WHITE, J. G.; SOUTHGATE, E.; THOMSON, J. N.; BRENNER, S. (1986). «The structure of the nervous system *C. elegans*». *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 314: 1-340.
- WOOD, W. B. (1988a). «Introduction to *C. elegans* biology». A: WOOD, W. B. [ed.]. *The Nematode Caenorhabditis elegans*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 1-16.
- (1988b). *The Nematode Caenorhabditis elegans*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.