Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III

Pierre Gönczy*†‡, Christophe Echeverri*†‡, Karen Oegema*†, Alan Coulson§, Steven J. M. Jones‖, Richard R. Copley†, John Duperon*, Jeff Oegema*, Michael Brehm*‡, Etienne Cassin*, Eva Hannak*, Matthew Kirkham*, Silke Pichler*†, Kathrin Flohrs*, Anoesjka Goessen*, Sebastian Leidel*, Anne-Marie Alleaume*‡, Cécilie Martin*‡, Nurhan Özlü*, Peer Bork† & Anthony A. Hyman*†

** Max-Planck-Institute for Cell Biology and Genetics (MPI-CBG), Pfotenhauerstrasse 108, D-01307 Dresden, Germany*
† European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, D-69117 Heidelberg, Germany
§ The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK
‖ Genome Sequence Centre, British Columbia Cancer Research Centre, 600 West 10th Avenue, Vancouver, British Columbia, V5Z-4E6, Canada
‡ Present addresses: Swiss Institute for Experimental Cancer Research (ISREC), CH-1066 Epalinges/Lausanne, Switzerland (P.G); Cenix BioScience GmbH, Pfotenhauerstrasse 108, D-01307 Dresden, Germany (C.E., M.B., A.-M.A., C.M.).

Genome sequencing projects generate a wealth of information; however, the ultimate goal of such projects is to accelerate the identification of the biological function of genes. This creates a need for comprehensive studies to fill the gap between sequence and function. Here we report the results of a functional genomic screen to identify genes required for cell division in *Caenorhabditis elegans*. We inhibited the expression of ~96% of the ~2,300 predicted open reading frames on chromosome III using RNA-mediated interference (RNAi). By using an *in vivo* time-lapse differential interference contrast microscopy assay, we identified 133 genes (~6%) necessary for distinct cellular processes in early embryos. Our results indicate that these genes represent most of the genes on chromosome III that are required for proper cell division in *C. elegans* embryos. The complete data set, including sample time-lapse recordings, has been deposited in an open access database. We found that ~47% of the genes associated with a differential interference contrast phenotype have clear orthologues in other eukaryotes, indicating that this screen provides putative gene functions for other species as well.

Genome sequencing projects identify large numbers of genes in the search for biological function. Function can be tested directly in the organism of interest by using an appropriate assay or inferred on the basis of knowledge gained from the study of related genes in other organisms. However, a significant fraction of genes uncovered by sequencing projects are 'new' and cannot be rapidly assigned functions by either method. Thus, there is a need for large-scale studies to test directly the function of new genes, and validate gene functions inferred from studies of homologous proteins in other organisms. Large-scale functional genomic studies have already been successful in *S. cerevisiae* (ref. 1; http://www.mips.biochem.mpg.de/proj/eurofan/index.html; http::://genome-www.stanford.edu/cgi-bin/sgd/search), but comparable analysis is still lacking in metazoans. Therefore, one of the challenges facing biologists today is to design high-throughput assays to assign cellular functions to genes emerging from metazoan sequencing projects.

Designing a large-scale screen for cell division genes

We sought to use a functional genomic approach in the early *C. elegans* embryo to identify a large set of genes necessary for cell-division processes, for the following reasons. First, the genome is sequenced, offering the possibility to conduct a comprehensive analysis[2]. Second, cell-division processes can be examined with high spatial and temporal resolution using time-lapse differential interference contrast (DIC) microscopy, allowing detection of even subtle deviations from the wild-type sequence of events[3]. Third, using RNA-mediated interference (RNAi), expression of a given gene in the embryo can be abolished in a sequence-specific manner by subjecting parental germ cells to corresponding double-stranded RNA (dsRNA)[4]. Notably, RNAi can be also applied to genes essential for cell division, as no mitoses occur between the time germ cells are subjected to dsRNA and fertilization. In contrast, although the targeted disruption of the ~6,200 open reading frames (ORFs) of

Table 1 Positive controls

	Phenotype
Undiluted	13/13 (100%)
1:1 dilution	12/13 (~92%)
1:3 dilution	6/13 (~46%)
1:7 dilution	4/13 (~31%)

Double-stranded RNA corresponding to 13 genes whose phenotype in the early embryo is known from mutational analysis were generated and injected, either undiluted or diluted (1:1, 1:3, 1:7) with a mixture of 3–7 different dsRNA species corresponding to genes not associated with phenotypes. The resulting embryos were analysed by time-lapse DIC microscopy; phenotypes were scored as described in Methods. The fraction of dsRNAs giving rise to the expected phenotype is given for each dilution.
The following 13 genes were tested *cyk-1* (ref. 14), *let-99* (ref. 35), *mei-1* (ref. 36), *nmy-2* (ref. 37), *ooc-3* (ref. 38), *par-1* (ref. 39), *par-2* (ref. 18), *par-3* (ref. 17), *par-5* (K. Kemphues, personal communication), *par-6* (ref. 40), *zen-4* (refs 41, 42), *zyg-9* (ref. 43), *zyg-11* (ref. 44).

Table 2 Overall outcome of screen

Distinct genes tested	2,174
Total with phenotype	281 (~12.9%)
DIC phenotypes	133 (~6.1%)
DIC + EL	125*
DIC + L/A	5
DIC only	3†
Progeny phenotypes (only)	139 (~6.4%)
EL only	78
L/A only	61
L1/L2	33
L3/L4	15‡
Adult	13
F0 sterile	9 (~0.4%)

EL, embryonic lethal; L/A, larval/adult phenotype; L1/L2, early larval defect; L3/L4, late larval defect; Adult, adult defect. See Methods, Table 3 and database (http://mpi-web.embl-heidelberg.de/dbScreen/) for more information.
* In some cases, embryonic lethality was partial only; progeny test not conducted for *par-2* (F58B6.3), but assumed to be embryonic lethal.
† For one of these dsRNAs (Y79H2A.11/Y75B8.36), the progeny test became embryonic lethal if conducted 36 h after injection (instead of the usual 24 h).
‡ For one of these dsRNAs (ZK1236.3), larvae died as L3/L4.

the *Saccharomyces cerevisiae* genome is yielding functional information at an unprecedented scale (http://www.mips.biochem.mpg.de/proj/eurofan/index.html; http::://genome-www.stanford.edu/cgi-bin/sgd/search), the detailed cellular function of essential genes cannot be examined in that paradigm.

RNAi has been used previously as a reverse genetic tool to probe the function of individual genes, or undertake small-scale screening of a complementary DNA library[5]. We wanted to expand

Table 3 133 genes on chromosome III associated with DIC phenotypes

A. Meiotic divisions and early post-fertilization events

A1. Progress through meiotic divisions (11 genes)

Male and female pronuclei do not become visible; embryos seem arrested during meiotic divisions.

Gene	Brief identification
C02F5.9	proteasome component C5
C23G10.4	19S proteasome regulatory particle subunit
C30C11.2	26S proteasome regulatory subunit S3
F23F12.6	26S proteasome regulatory complex subunit p48A
F25B5.4	polyubiquitin (*ubq-1*)
F35G12.9	RING-finger containing protein (APC11-like)
F54C8.3	APC4 homologue (*emb-30*) *
F57B9.10	26S proteasome regulatory complex subunit p42B
K06H7.5	APC subunit 2
T05G5.3	Cdk1/Cdc2 kinase (*ncc-1*) *
ZK1010.1	ubiquitin (*ubq-* 2)

A2. Fidelity of meiotic divisions (27 genes)

Multiple female pronuclei; irregular cytoplasm; aberrant pseudoclevage stage; spindle unstable during anaphase; karyomeres in AB/ P1; AB/P1 nuclei off-centre; often semi-sterile.

B0336.10	60S ribosomal protein L17A
B0393.1	40S ribosomal protein SA
B0412.4	40S ribosomal protein S29
B0464.1	aspartyl-tRNA synthetase
C14B9.7	ribosomal protein L21
C16A3.9	40S ribosomal protein S13
C23G10.3	ribosomal protein S3
C27D11.1	translation initiation factor 3 subunit 10 (*egl-45*)
C54C6.1	60S ribosomal protein L37
F13B10.2	60S ribosomal protein L3
F26F4.10	arginyl tRNA synthetase
F37C12.4	ribosomal protein YL39
F37C12.9	ribosomal protein S14
F37C12.11	ribosomal protein S21
F53A3.3	40S ribosomal protein
F54E7.2	ribosomal protein S12
F56F3.5	ribosomal protein S3a
F57B9.3	eukaryotic initiation factor 4A
H06I04.4	ubiquitin-like ribosomal protein S27A fusion protein
R08D7.3	translation initiation factor eIF3 p66 subunit
R13A5.8	ribosomal protein L9
R74.1	leucyl-tRNA synthetase
R151.3	ribosomal protein ML16
T05G5.10	translation initiation factor eIF5A
T10F2.1	glycyl-tRNA synthetase
T20H4.3	prolyl-tRNA synthetase region
ZK652.4	60S ribosomal protein L35

A3. Entry into interphase (2 genes)

Delay before entering interphase; vigorous cytoplasmic and cortical movements; aberrant number and/or position of pronuclei.

F48E8.5	protein phosphatase 2A regulatory subunit
ZK520.4	cullin (*cul-2*) *

A4. Pseudocleavage stage (2 genes)

Little/no cortical ruffing or pseudocleavage furrow.

C34C12.3	serine/threonine protein phosphatase
Y49E10.19	homologies to anilin (actin-binding protein)

B. Nuclear appearance

B1. Pronuclear/nuclear appearance (5 genes)

Pronuclei and nuclei in daughter blastomeres are not/poorly visible; spindle not/poorly visible; often failure in cytokinesis

C29E4.3	Ran-GAP1
C38D4.3	weak homologies to calpastatin
F59A2.1	Ran-binding protein 2 (NUP 358)
K01G5.4	Ran
ZK328.5	NUP 98 (nucleoporin)

B2. Nuclear appearance (5 genes)

Nuclei in daughter blastomeres are not/poorly visible

C29E4.2	no clear homologies
F34D10.2	strong homologies to CDC45
R10E4.4	DNA replication licensing factor MCM5
Y55B1BR.3	contains chromo domain
ZK632.1	DNA replication licensing factor MCM6

C. Cell division processes in the early embryo

C1. Pronuclear migration (6 genes)

Lack of male pronuclear migration; female pronuclear migration variable; sometimes multiple female pronuclei; no or small spindle.

C05D11.3	putative ATP binding protein
C28H8.12	dynamitin (p50) (*dnc-2*) *
C36E8.5	beta-tubulin
K01G5.7	beta-tubulin
T03F6.5	LIS-1 homologue
T26A5.9	dynein 8kd light chain

C2. Spindle assembly (2 genes)

Spindle is very small or not visible; karyomeres are generated

F58A4.8	gamma-tubulin
H04J21.3	no clear homologies

C3. Fidelity of mitotic divisions (cross-eye phenotype) (3 genes)

Daughter nuclei stay close to central cortex; usually karyomeres in daughter blastomeres.

C02F5.1	homologies to coiled coil containing proteins
F35G12.8	SMC-1 condensin
F58A4.3	histone H3-like protein (homology to CENP-A)

C4. Fidelity of mitotic divisions (karyomeres only) (3 genes)

Karyomeres in daughter blastomere AB and/or P1.

F54C8.2	histone H3-like protein (homology to CENP-A)
R107.6	homologous to *Drosophila orbit*
Y43F4B.6	kinesin (KIF4-like)

C5. Anaphase spindle positioning (symmetric division) (4 genes)

No posterior spindle displacement during anaphase; symmetric division.

C38C10.4	no clear homologies outside *C. elegans*
F22B7.13	no clear homologies outside *C. elegans*
F54E7.3	PDZ containing protein PAR-3 (*par-3*) *
F58B6.3	RING-finger containing protein PAR-2 (*par-2*) *

C6. Cytokinesis (5 genes)

Cleavage furrow not visible or regresses.

B0464.5	serine/threonine kinase; similar to *S. pombe* DSK-1
C56G7.1	nonmuscle myosin regulatory light chain (*mlc-4*) *
F11H8.4	formin homology protein CYK-1 (*cyk-1*) *
K08E3.6	Rho family GTPase activating protein (*cyk-4*) *
T25C8.2	actin (*act-5*)

C7. P1 rotation (2 genes)

No rotation of centrosome/nuclear complex in P1.

F55H2.3	*let-99* homologue
Y39E4B.1	RNAse L inhibitor

D. Pace of development

D1. General pace of development (overall slow) (12 genes)

Slow overall pace of development (>30 min between pronuclear migration and AB division -compared to 18-22 min in wt).

C29E4.8	adenylate kinase
C34E10.6	ATP synthase beta chain
F23H11.5	no clear homologies
F35G12.2	isocitrate dehydrogenase
F35G12.10	ATP synthase B chain
F54H12.1	aconitate hydratase
F56D2.1	mitochondrial processing protease enhancing protein
K04G7.4	NADH dehydrogenase
T07C4.7	succinate dehydrogenase cytochrome b chain (*mev-*
T20G5.2	citrate synthase
T27E9.1	ADP/ATP mitochondrial carrier protein
ZK637.8	vacuolar H+-ATPase (TJ6/proton pump) (*unc-32*)

D2. Pace of development (8 genes)

Slow between pseudocleavage stage and pronuclear envelope breakdown; P1 division delayed with respect to that of AB.

C03C10.3	ribonucleotide reductase small subunit
F31E3.3	replication factor C complex protein
F44B9.7	replication factor C subunit 3
F58A4.4	DNA primase 49Kd subunit
R01H10.1	DNA polymerase alpha/primase complex chain B
T23G5.1	ribonucleotide-disphosphate reductase large chain
T24C4.5	DNA primase subunit
Y47D3A.29	DNA polymerase alpha-subunit

E. Embryo appearance and morphology

E1. Osmotic integrity and other processes (16 genes)

Embyros loose structural integrity upon dissection ; limited phenotypic analysis done in utero.

B0336.2	ADP ribosylation factor 1 (*arf-1*)
B0361.10	glycosyl transferase group 1 protein
C07H6.5	RNA helicase (DEAD-box protein family)
C36A4.4	UDP-N-acetylglucosamin
D2045.1	no clear homologies
F08F8.2	3-hydroxy-3-methylglutaryl-Coenzyme A reductase
F55H2.2	probable vacuolar ATP synthase subunit D
K10D2.6	NADPH-cytochrome P450
K12H4.4	signal peptidase subunit
PAR2.4	homologies to proteins in other metazoans
T05G5.7	no clear homologies outside *C. elegans*
T12A2.2	oligosaccharyl transferase STT3 subunit
ZK328.1	ubiquitin carboxyl-terminal hydrolase
ZK512.5	homologies to *Drosophila* and human proteins
ZK686.3	homologies to oligosaccharyl transferase 34 kd subunit
Y76A2B.1	coronin-like protein (*pod-1*) *

E2. Cytoplasmic appearance (sparse yolk granules) (4 genes)

Markedly reduced density of yolk granules throughout embryo.

C45G9.5	no clear homologies
T20G5.1	clathrin heavy chain
ZK1098.5	putative secretory protein (Bet3p-like)
Y37D8A.10	homologies to signal peptidase complex 25 kDa

E3. Cytoplasmic appearance (irregular) (3 genes)

Uneven or irregular distribution of yolk granules.

C03C10.1	casein kinase I
T20B12.1	homologies to O-linked GlcNAc transferase
Y49E10.15	small nuclear ribonucleoprotein E (snRNP E)

F. Unique phenotypes (13 genes)

Phenotypes associated with single genes: see comprehensive Table 3 in supplementary information for specifics.

C07G2.3	T-complex protein 1, epsilon subunit (*cct-5*)
C14B9.4	polo-like kinase (*plk-1*) *
C18D11.5	homologies with transcription factors
F01F1.8	T-complex protein 1, zeta subunit (*cct-6*)
F10E9.8	no clear homologies
H38K22.2	homologies to proteins in other eukaryotes
K11D9.1	kinesin CeMCAK (KIF2/XKCM1-like)
R151.9	homologies to c-myc binding protein MM-1
T04A8.7	1,4-alpha-glucan branching enzyme
Y79H2A.11	doublecortin-related kinase
Y41C4A.10	Elongin B (regulator of RNA polymerase II)
Y49E10.1	26S proteasome regulatory subunit 8
Y56A3A.20	CCR4-associated factor 1

The gene name and a brief protein identification are given, as is a brief description of the phenotype characteristic of each class. Asterisk indicates that the phenotype for this gene was previously reported. Phenotypes represented by a single gene (two in the case of *cct-5* and *cct-6*) are grouped in 'unique phenotypes' for simplicity; specifics about each unique phenotype can be found in the Supplementary Information. A comprehensive version of Table 3 is available in the Supplementary Information. Additional details and time-lapse recordings can be found at http://mpi-web.embl-heidelberg.de/dbScreen/.

on these studies and apply RNAi on a genomic scale to identify genes required for cell division. We therefore sought to devise a screen that allows unbiased testing of each predicted ORF, not only of those represented in cDNA libraries. Here we report the design of such a screen, and its application to test the function of 2,232 (~96%) of the predicted ORFs on chromosome III of *C. elegans*.

Establishing an efficient screening paradigm

In brief, the screening strategy was established as follows. Primer pairs corresponding to individual ORFs were selected using a customized algorithm, with addition of T7 and T3 promoter sequences. Corresponding pieces of wild-type genomic DNA were PCR-amplified, separate T3 and T7 RNA transcription reactions performed and the resulting single-stranded RNAs (ssRNAs) annealed to generate dsRNA. After verifying their quality by gel electrophoresis, dsRNAs were micro-injected into wild-type hermaphrodites, and the resulting embryos were analysed 24 h later using time-lapse DIC microscopy. Single embryos were filmed from shortly after fertilization until the four-cell stage. This DIC assay constituted the core of the screen, and was optimized for detecting even minor deviation from the wild-type sequence of events. At the time of filming, three injected animals were also transferred to a fresh plate, which was scored 2 and 4 days later using low-magnification stereomicroscopy for the presence of progeny, and for any gross deviation from wild-type development.

To test whether this screening strategy would lead to the efficient identification of genes required in the early embryo, we injected individual dsRNAs corresponding to 13 genes characterized previously by mutational analysis to give rise to a phenotype visible by time-lapse DIC microscopy. As reported in Table 1, the expected phenotype was identified in all 13 cases. To determine the pool size optimal for large-scale studies, each of the 13 dsRNAs was mixed with 1, 3 or 7 volumes of dsRNA corresponding to genes not associated with phenotypes. Such dilutions substantially dampened the potency of RNAi (Table 1). Dampening was not observed when dsRNAs were diluted with either injection buffer or ssRNA (data not shown). These results suggest that there is a rapidly titratable step necessary for dsRNAs to generate RNAi. On the basis of these results, we chose a pool size of two to achieve higher throughput but still ensure identification of over 90% of the predicted genes required in the early embryo.

Pairs of randomly chosen dsRNAs were injected, and the resulting progeny were analysed with both DIC assay and progeny test. Pairs giving rise to a phenotype were split, and the two individual dsRNAs tested in turn. Each dsRNA associated with a phenotype was analysed further to ascertain that the phenotype did result from inactivation of the expected gene. This was necessary because of two potential caveats. First, individual dsRNAs can abolish expression of several genes closely related at the nucleotide level[6,7]. Therefore, a BlastN homology search was performed with the PCR-amplified piece of DNA against the *C. elegans* genome. When this revealed sequences with more than 80% nucleotide identity over 200 base pairs (bp), or more than 90% over 100 bp, new primers were designed wherever possible to avoid regions of high sequence similarity. Second, mispredicted gene boundaries may result in some dsRNAs that also inactivate neighbouring genes. Therefore, dsRNAs whose neighbour was also associated with a phenotype were identified. When the neighbour was predicted to be less than 1 kb away from the amplified DNA, new primer pairs were designed wherever possible to increase that distance to more than 1 kb. In addition to these two steps, an independent dsRNA was generated for most genes associated with a DIC phenotype to verify the validity of the initial result (see Supplementary Information). Those genes whose RNAi phenotype was lost during this confirmation round were excluded from the data set. As a result, the DIC phenotypes reported here almost certainly result from inactivation of the expected genes.

Analysis of ~96% of genes on chromosome III

We tested 2,232 of the 2,315 predicted ORFs (96.4%) on chromosome III and found that dsRNAs corresponding to 133 genes (6.1% of distinct genes tested) gave rise to a phenotype detectable in the DIC assay (Tables 2 and 3). dsRNAs corresponding to 139 genes gave rise to a phenotype detectable solely in the progeny test: 78 genes were associated with embryonic lethality, 48 with a larval phenotype; 13 with an adult phenotype; in addition, 9 dsRNAs resulted in sterility of the injected animals (Table 2; and Supplementary Information Table 4).

We compared these results to those obtained with classical genetics (see Supplementary Information). This showed that we had identified all seven chromosome III genes known to mutate to a phenotype recognizable by DIC microscopy in the early embryo, and the four genes whose role in the early embryo had been previously unravelled using RNAi[5,8–18]. Moreover, we identified 9 out of 14 (~64%) loci known to mutate to phenotypes detectable later during embryogenesis, and 9 out of 31 (~29%) loci known to mutate to phenotypes in larvae or adults and that could have been easily recognized in the progeny test. These observations are compatible with a previous report that RNAi is less potent in silencing gene expression later during development[4]. Longer exposure to dsRNA, after injection or feeding, may help circumvent this limitation.

Notably, this analysis suggests that we have identified the large majority of the genes on chromosome III that are required for cellular processes scoreable by DIC microscopy in the early embryo. As we tested ~12% of the predicted ORFs in the genome, and assuming a random genome-wide distribution, it can be inferred that over 1,000 genes are essential for the first two cleavage divisions of *C. elegans*.

RNAi as an efficient gene function discovery tool

Detailed analysis of time-lapse recordings allowed us to place the 133 genes associated with a DIC phenotype into distinct phenotypic classes (Table 3). Owing to space limitations, only some classes are discussed here and illustrated in Fig. 1. The complete data set, along with time-lapse DIC recordings, can be found on an open-access database (http://mpi-web.embl-heidelberg.de/dbScreen/).

Of the 133 genes associated with a DIC phenotype, only 11 (8.2%) had been previously ascribed a function by direct experimentation in the early *C. elegans* embryo[5,8–18]. Another 104 (78.2%) have homologues in other species, from which a function in *C. elegans* might potentially be derived. For the remaining 18 (13.5%), sequence information alone did not yield obvious clues to a possible function. Therefore, this screen represents an efficient discovery tool to assign gene function in the cognate cellular context. We illustrate this point with three phenotypic classes.

In the 'pronuclear/nuclear appearance' phenotypic class (B1 in Table 3), comprising five genes, pronuclei in the one-cell stage embryo, as well as nuclei in daughter blastomeres, are poorly or not visible (Fig. 1b). In addition, the spindle is poorly or not visible, and the cleavage furrow often regresses, probably as a consequence of the spindle defect. We infer that these genes are required for nuclear envelope assembly and an aspect of spindle formation. Notably, three of these genes encode the small GTPase Ran and its activating proteins Ran-GAP1 and Ran-BP2, which have a crucial role in nuclear envelope assembly and spindle formation in *Xenopus* egg extracts[19–24]. Our work confirms that these molecules have a similar role *in vivo*. A fourth gene in this phenotypic class encodes the *C. elegans* orthologue of the nucleoporin NUP 98, and a fifth one a protein with weak homologies to calpastatin. An attractive hypothesis is that NUP 98 and a calpastatin-related protein may also

participate in nuclear envelope assembly in *C. elegans* and perhaps other metazoans.

In the 'pronuclear migration' phenotypic class (C1 in Table 3), comprising six genes, migration of pronuclei is affected (Fig. 1d). Two of the corresponding genes encode β-tubulin subunits, an expected finding as pronuclear migration is a microtubule-dependent process[25]. Two additional genes encode a dynein light chain and the dynactin component p50/dynamitin, which was also anticipated because cytoplasmic dynein function is required for this process in *C. elegans*[9]. In contrast, the involvement of the

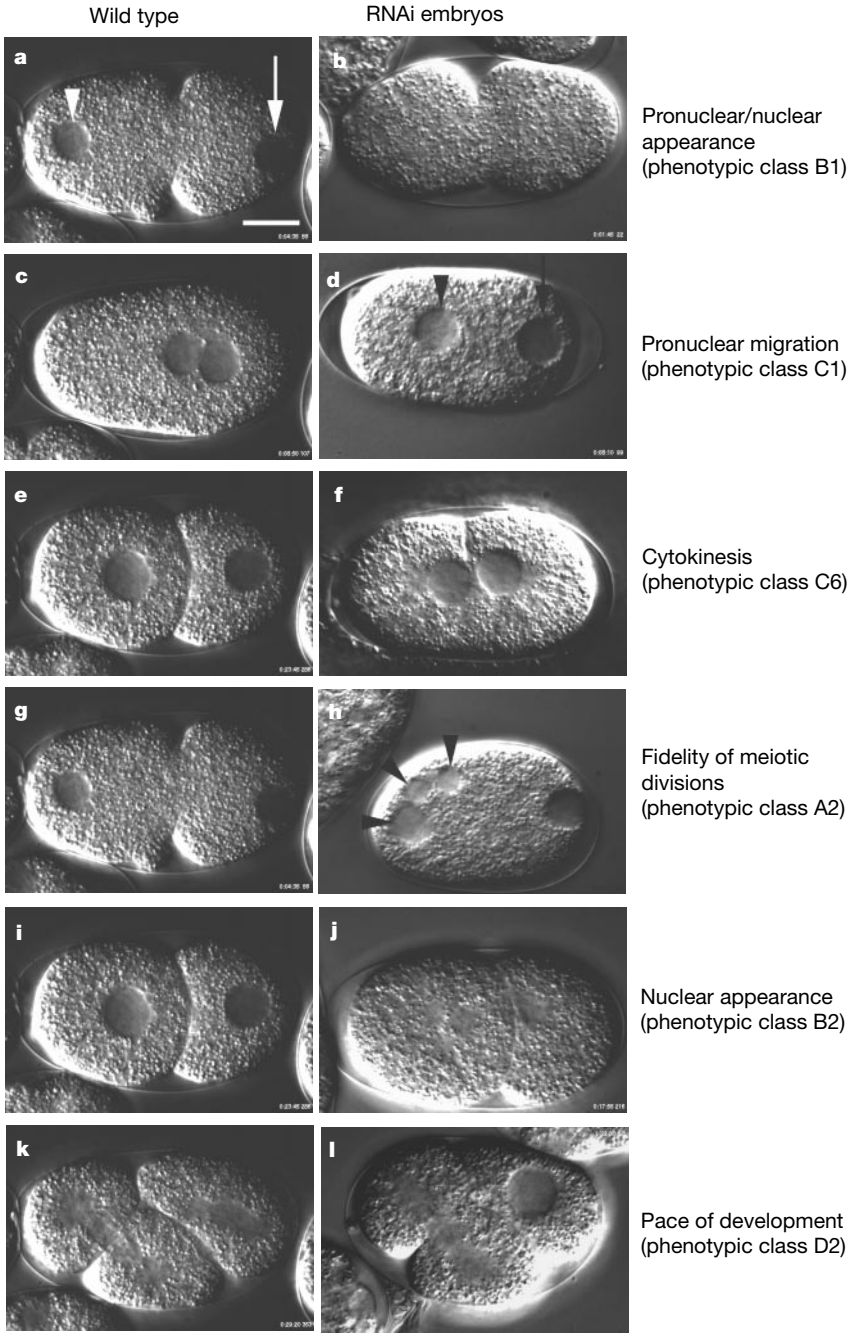

Figure 1 Single images taken from time-lapse DIC recordings of wild-type (left) and embryos representative of six distinct phenotypic classes (right). Anterior is to the left and posterior to the right. All panels at same magnification; scale bar,10 μm.
a, **b**, Pseudocleavage stage. In wild type (**a**), both male (arrow) and female (arrowhead) pronuclei are visible as clear discs excluding yolk granules. In contrast, in embryos of the pronuclear/nuclear appearance class (**b**; C29E4.3, RanGap1 shown here), pronuclei are not visible. Note presence of pseudocleavage furrow in both cases. **c**, **d**, Pronuclear migration stage. In wild type (**c**), the two pronuclei have met in the posterior half of the embryo. In embryos of the pronuclear migration class (**d**; T03G6.5, LIS-1 shown here), the male pronucleus (arrow) is still in the vicinity of the posterior cortex, the female pronucleus (arrowhead) is in the anterior of the embryo. **e**–**f**, Two-cell stage. In wild-type (**e**), blastomeres of the two-cell stage embryo have been clearly separated by the cleavage furrow. In embryos of the cytokinesis class (**f**; B0464.5, serine/threonine kinase shown here), the cleavage furrow regresses, and daughter nuclei rejoin in a single cell. **g**, **h**, Pseudocleavage stage. In wild-type (**g**), there is a single female pronucleus. In embryos of the fidelity of meiotic divisions class (**h**; C16A3.9, 40S ribosomal protein S13 shown here), there are several female pronuclei (arrowheads). Note absence of pseudocleavage furrow. **i**, **j**, Two-cell stage. In wild-type (**i**), nuclei in daughter blastomeres are visible as clear discs excluding yolk granules. In embryos of the nuclear appearance class (**j**; F34D10.2, CDC45-like shown here), nuclei are not visible. **k**, **l**, About 20 min after pronuclear migration. In wild type (**k**), the anterior blastomere (AB) is completing division, whereas the posterior blastomere (P_1), follows 2–3 min later; note that the nuclear envelope in P_1 has already broken down. In embryos of the pace of development class (**l**; T24C4.5, DNA primase shown here), division of P_1 is delayed with respect to that of AB; note intact nuclear envelope in P_1.

remaining two genes was not documented previously in nematodes. We found that a putative ATP-binding protein and the worm homologue of lis-1, the gene mutated in Müller–Dieker lissencephaly, are also required for pronuclear migration. lis-1 homologues act in concert with dynein to mediate nuclear localization in both *Aspergillus* and *Drosophila*[26,27], and our finding indicates that the two components may also function together in *C. elegans*.

In the 'cytokinesis' phenotypic class (C6 in Table 3), comprising five genes, cleavage is defective, resulting in the rejoining of daugther blastomeres (Fig. 1f). Three of these genes, *cyk-1*, *cyk-4* and *mlc-4*, which encode a formin[14], a Rho GAP[15] and a non-muscle myosin regulatory light chain[5], respectively, have been initially identified by mutational analysis. This confirms that an RNAi-based functional genomic approach will provide a rich source for the molecular identification of loci defined by classical genetics.

Functional groups of genes required for cellular processes

Our RNAi-based functional genomic screen has also allowed us to reach unexpected conclusions about the involvement of particular biochemical pathways in specific cellular process in the early *C. elegans* embryo. We illustrate this point with three phenotypic classes.

In the 'fidelity of meiotic divisions' class (A2 in Table 3), comprising 27 genes, embryos often display several female pronuclei, indicative of defects during the female meiotic divisions (Fig. 1h). Additional phenotypic manifestations include an unstable spindle during anaphase and generation of karyomeres in daughter blastomeres. Unexpectedly, all 27 genes in this class encode components of the translation machinery, suggesting that efficient protein synthesis is required for proper execution of meiotic divisions in *C. elegans*. Whether this reflects a requirement for overall protein synthesis, or a requirement for the synthesis of a specific subset of proteins needed for the meiotic divisions remains to be determined.

In the 'nuclear appearance' phenotypic class (B2 in Table 3), comprising five genes, although pronuclei are visible as in wild-type, nuclei in daughter blastomeres are poorly or not visible (Fig. 1j). We infer that these genes are required specifically for the reassembly of the nucleus after mitosis. Three of these genes encode *C. elegans* orthologues of the DNA replication licensing factors CDC45, CDC46 and MCM2/3, which are required both for DNA replication and for restricting its occurrence to once per cell cycle[28]. These findings suggest that there may possibly be a mechanism that monitors the presence of DNA licensing factors before nuclear reassembly can be triggered.

In the 'pace of development' phenotypic class (D2 in Table 3), comprising eight genes, events are delayed in the one-cell stage between pronuclear meeting and pronuclear envelope breakdown, and in the two-cell stage most strikingly in the P_1 blastomere (Fig. 1l). Notably, six of these eight genes encode components of the DNA replication machinery, whereas the remaining two encode polypeptides of ribonucleotide-diphosphate reductase, which is required to produce deoxyribonucleotides. This indicates that the observed delays may be due to defects in DNA replication. This implicates the existence of an hitherto unsuspected DNA replication checkpoint in the early *C. elegans* embryo. This checkpoint mechanism may act preferentially in the P lineage, as events appear less delayed in the AB blastomere of the two-cell stage embryo.

Toward comprehensive metazoan functional genomics

We wanted to investigate whether genes identified in this screen are likely to have similar functions in other organisms. As an indirect way to address this issue, we determined whether genes associated with DIC phenotypes were more likely to be conserved across evolution. We classified all tested proteins into one of four groups based on BlastP analysis: (1) nematode-specific; (2) metazoan-specific; (3) eukaryotic-specific; (4) multi-kingdom. We found a distinct enrichment of evolutionarily conserved genes among those associated with DIC phenotypes (Fig. 2a). Most strikingly, although *C. elegans* genes that have orthologues in both *Drosophila* and *S. cerevisiae* represent only 12.9% of genes tested, they comprise 47.3% of those associated with DIC phenotypes (Fig. 2b). Therefore, genes required for proper division of the early *C. elegans* embryo tend to have close relatives in other species, suggesting that a number of the functional assignments made here may be transferred to other eukaryotes.

Our work shows that RNAi is an efficient reverse genetic tool to undertake a large-scale functional genomic analysis. Detailed analysis using time-lapse DIC microscopy has allowed us to identify 'phenotypic signatures' characteristic of functional groups of genes, and thus establish a unique functional database of cell-division proteins. By continuing this analysis on a genome-wide basis, by using RNAi to probe other processes with different assays, and by comparing the data with functional information emerging from other organisms, a significant fraction of metazoan genes may be assigned functions. □

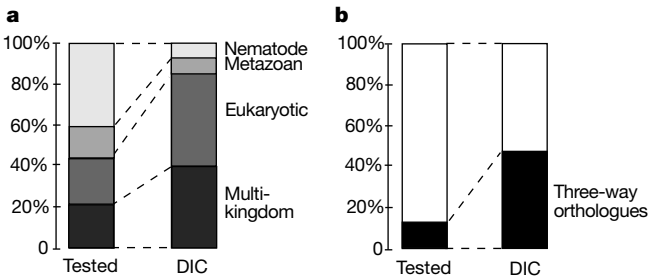

Figure 2 Distribution of predicted proteins according to homology and orthology relationships. Left, 2,233 genes tested (2,232 from Wormpep20 plus Y47D3A.29); right, 133 genes associated with DIC phenotypes. **a**, Homology. BlastP analysis was conducted to place predicted proteins into one of the following categories (from top to bottom, in increasing shadings): nematode-specific (39.9% in tested set, 7.5% in DIC set); metazoan-specific (16.3% in tested set, 7.5% in DIC set); eukaryotic-specific (23% in tested set, 45.1% in DIC set); multi-kingdom (20.7% in tested set, 39.8% in DIC set). **b**, Three-way (*C. elegans*, *Drosophila*, *S. cerevisiae*) reciprocal BlastP analyses were conducted to determine whether predicted proteins had clear orthologues in other eukaryotes (black shading: 12.9% in tested set; 47.3% in DIC set).

Methods

Selection of primer pairs and number of genes tested

Primer pairs were selected using a customized program to amplify the shortest possible region containing > 500 bp or >90% coding sequence, with preference given to experimentally confirmed exons and regions covered by expressed sequence tags (ESTs) (see Supplementary Information).

We tested 2,232/2,315 chromosome III ORFs (Wormpep20) (see list at http://mpi-web.embl-heidelberg.de/dbScreen/). These correspond to 2,174 distinct genes because (1) 48 genes have two or more alternatively spliced variants; (2) F55H2.3/F55H2.4, K06H7.5/K06H7.6, Y79H2A.11/Y75B8A.36 each correspond to a single gene (Table 3); and (3) one gene associated with a phenotype (Y47D3A.29) was not present in Wormpep20.

Generation of dsRNA

All steps were done in 96-well plates. PCR reactions (50 μl) were performed using 0.8 μM primers and ~0.1 μg wild-type template genomic DNA. PCR products were ethanol-precipitated and resuspended in 7.0 μl TE buffer; 1.0 μl was used for separate 5 μl transcription reactions using T3 and T7 RNA polymerases (Ambion). The reactions were diluted to 50 μl with RNase-free water, combined, and then the mixed RNA was purified (Qiagen) into 130 μl RNase-free H_2O; 50 μl of this was mixed with 10 μl 6× injection buffer[29], and the RNA annealed by heating for 10 min at 68 °C, and 30 min at 37 °C. Concentrations of dsRNAs were 0.1–0.3 μg μl^{-1}. Double stranding was assessed by scoring shifts in mobility with respect to ssRNA on 1% agarose gels.

Injections and time-lapse DIC microscopy assay

Pairs of dsRNAs were injected bilaterally into the gonads of adult wild-type hermaphrodites[29], which were left at 20 °C for 24 h. Embryos were then removed and

analysed for potential defects in cell-division processes, capturing 1 image every 5 s (ref. 3). Details about cellular processes scored are given in Supplementary Information. A minimum of three embryos from three different worms were filmed from shortly after fertilization until the four-cell stage. Embryos from two additional worms were also inspected and still images taken. Embryos resulting from injection of some dsRNAs lost structural integrity on dissection, and were examined *in utero*[3].

Pairs of dsRNAs were split if one or more embryo displayed a phenotype. A phenotype was attributed to a given gene when it was observed in at least two embryos examined after injection of individual dsRNA.

Progeny test

Three animals were transferred to a fresh plate 24 h after injection, and left at 20 °C. Two days later, the plate was inspected with a stereomicroscope (×20–40 magnification) for the presence of eggs, F1 larvae and their developmental stage (normally L2–L4). Two days after that, the plate was inspected for the presence of F1 adults (normally >100), their overall body morphology and the presence of F2 progeny. Whenever the progeny test was conducted more than once with a given dsRNA and gave rise to different results, the strongest phenotype was retained.

Double-stranded RNAs that gave rise to eggs and 0–10 F1 larvae were deemed embryonic lethal. dsRNAs that gave rise to a phenotype detectable in larvae fell into two groups: those in which the progeny on day four resembled L1 or L2 ('early larval defect'); and those in which the progeny on day four resembled L3 or L4 ('late larval defect'). Partially penetrant embryonic lethality and subtle developmental defects were not scored in this analysis. Moreover, dsRNAs that gave rise to defects in less than 5% of the adult progeny were not considered as being associated with a phenotype.

Bioinformatic analysis

Homologues. *C. elegans* protein sequences were searched against the sp_nrdb database, a non-redundant composite of SWISS-PROT and TrEMBL[30], using BlastP[31] and masking compositionally biased regions with the SEG filter. Significantly scoring sequences (E-value $< 10^{-6}$) were retrieved, and species and phyla information extracted from the sequence entry file. On the basis of these results, proteins were broken down into the following categories: (1) nematode specific (no significant homologue present outside nematodes); (2) metazoan specific (significant homologues present in other metazoans, but not outside metazoans); (3) eukaryotic specific (significant homologues present among unicellular eukaryotes, but not outside eukaryotes); (4) multi-kingdom (significant homologues present in both eukaryotes and prokaryotes). Although this procedure will not recognize very distant homologues, it will identify proteins that are well-conserved across kingdoms. Varying the E-value to $<10^{-3}$ did not lead to major changes in the outcome of the analysis.

Orthologues. Using BlastP[31], we identified proteins from *C. elegans*, *D. melanogaster* and *S. cerevisiae* that were all each other's best match in the respective genomes[32]. That two proteins in different genomes are each others host match is commonly used as an operational definition of orthology[33,34], although it is recognized that full reconstructions of phylogeny are required to resolve complex evolutionary histories. *C. elegans* protein data were obtained by embedding in Wormpep20 the protein sequences tested by RNAi. *Drosophila* protein sequence data were obtained from the Celera genome CD-ROM, *S. cerevisiae* protein data from the NCBI.

Received 17 July; accepted 17 October 2000.

1. Ross-Macdonald, P. *et al.* Large-scale analysis of the yeast genome by transposon tagging and gene disruption. *Nature* **402,** 413–418 (1999).
2. The *C. elegans* Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282,** 2012–2018 (1998).
3. Gönczy, P. *et al.* Dissection of cell division processes in the one cell stage *Caenorhabditis elegans* embryo by mutational analysis. *J. Cell Biol.* **144,** 927–946 (1999).
4. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391,** 806–811 (1998).
5. Shelton, C. A., Carter, J. C., Ellis, G. C. & Bowerman, B. The nonmuscle myosin regulatory light chain gene *mlc-4* is required for cytokinesis, anterior-posterior polarity, and body morphology during *Caenorhabditis elegans* embryogenesis. *J. Cell Biol.* **146,** 439–451 (1999).
6. Zhang, B. *et al.* A conserved RNA-binding protein that regulates sexual fates in the *C. elegans* hermaphrodite germ line. *Nature* **390,** 477–484 (1997).
7. Cox, D. N. *et al.* A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes Dev.* **12,** 3715–3727 (1998).
8. Skop, A. R. & White, J. G. The dynactin complex is required for cleavage plane specification in early *Caenorhabditis elegans* embryos. *Curr. Biol.* **8,** 1110–1116 (1998).
9. Gönczy, P., Pichler, S., Kirkham, M. & Hyman, A. A. Cytoplasmic dynein is required for distinct aspects of MTOC positioning, including centrosome separation, in the one cell stage *Caenorhabditis elegans* embryo. *J. Cell Biol.* **147,** 135–150 (1999).
10. Chase, D. *et al.* The polo-like kinase PLK-1 is required for nuclear envelope breakdown and the completion of meiosis in *Caenorhabditis elegans*. *Genesis* **26,** 26–41 (2000).
11. Boxem, M., Srinivasan, D. G. & van den Heuvel, S. The *Caenorhabditis elegans* gene *ncc-1* encodes a cdc2-related kinase required for M phase in meiotic and mitotic cell divisions, but not for S phase. *Development* **126,** 2227–2239 (1999).
12. Feng, H. *et al.* CUL-2 is required for the G1-to-S-phase transition and mitotic chromosome condensation in *Caenorhabditis elegans*. *Nature Cell Biol.* **1,** 486–492 (1999).
13. Rappleye, C. A., Paredez, A. R., Smith, C. W., McDonald, K. L. & Aroian, R. V. The coronin-like protein POD-1 is required for anterior-posterior axis formation and cellular architecture in the nematode *Caenorhabditis elegans*. *Genes Dev.* **13,** 2838–2851 (1999).
14. Swan, K. A. *et al.* cyk-1: a *C. elegans* FH gene required for a late step in embryonic cytokinesis. *J. Cell Sci.* **111,** 2017–2027 (1998).
15. Jantsch-Plunger, V. *et al.* CYK-4. A rho family gtpase activating protein (gap) required for central spindle formation and cytokinesis. *J. Cell Biol.* **149,** 1391–1404 (2000).
16. Furuta, T. *et al.* EMB-30: an APC4 homologue required for metaphase-to-anaphase transitions during meiosis and mitosis in *Caenorhabditis elegans*. *Mol. Biol. Cell* **11,** 1401–1419 (2000).
17. Etemad-Moghadam, B., Guo, S. & Kemphues, K. J. Asymmetrically distributed PAR-3 protein contributes to cell polarity and spindle alignement in early *C. elegans* embryos. *Cell* **83,** 743–752 (1995).
18. Levitan, D. J., Boyd, L., Mello, C. C., Kemphues, K. J. & Stinchcomb, D. T. *par-2*, a gene required for blastomere asymmetry in *Caenorhabditis elegans*, encodes zinc-finger and ATP-binding motifs. *Proc. Natl Acad. Sci. USA* **91,** 6108–6112 (1994).
19. Nicolas, F. *et al. Xenopus* Ran-binding protein 1: molecular interactions and effects on nuclear assembly in *Xenopus* egg extracts. *J. Cell Sci.* **110,** 3019–3030 (1997).
20. Carazo-Salas, R. E. *et al.* Generation of GTP-bound Ran by RCC1 is required for chromatin-induced mitotic spindle formation. *Nature* **400,** 178–181 (1999).
21. Kalab, P., Pu, R. T. & Dasso, M. The ran GTPase regulates mitotic spindle assembly. *Curr. Biol.* **9,** 481–484 (1999).
22. Wilde, A. & Zheng, Y. Stimulation of microtubule aster formation and spindle assembly by the small GTPase Ran. *Science* **284,** 1359–1362 (1999).
23. Hetzer, M., Bilbao-Cortes, D., Walther, T. C., Gruss, O. J. & Mattaj, I. W. GTP hydrolysis by Ran is required for nuclear envelope assembly. *Mol. Cell* **5,** 1013–1024 (2000).
24. Zhang, C. & Clarke, P. R. Chromatin-independent nuclear envelope assembly induced by Ran GTPase in *Xenopus* egg extracts. *Science* **288,** 1429–1432 (2000).
25. Strome, S. & Wood, W. B. Generation of asymmetry and segregation of germ-line granules in early *C. elegans* embryos. *Cell* **35,** 15–25 (1983).
26. Swan, A., Nguyen, T. & Suter, B. Drosophila Lissencephaly-1 functions with Bic-D and dynein in oocyte determination and nuclear positioning. *Nature Cell Biol.* **1,** 444–449 (1999).
27. Willins, D. A., Liu, B., Xiang, X. & Morris, N. R. Mutations in the heavy chain of cytoplasmic dynein suppress the nudF nuclear migration mutation of *Aspergillus nidulans*. *Mol. Gen. Genet.* **255,** 194–200 (1997).
28. Tye, B. K. MCM proteins in DNA replication. *Annu. Rev. Biochem.* **68,** 649–686 (1999).
29. Mello, C. C., Kramer, J. M., Stinchcomb, D. & Ambros, V. Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* **10,** 3959–3970 (1991).
30. Bairoch, A. & Apweiler, R. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* **28,** 45–48 (2000).
31. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–3402 (1997).
32. Tatusov, R. L., Koonin, E. V. & Lipman, D. J. A genomic perspective on protein families. *Science* **278,** 631–637 (1997).
33. Chervitz, S. A. *et al.* Comparison of the complete protein sets of worm and yeast: orthology and divergence. *Science* **282,** 2022–2028 (1998).
34. Snel, B., Bork, P. & Huynen, M. A. Genome phylogeny based on gene content. *Nature Genet.* **21,** 108–110 (1999).
35. Rose, L. S. & Kemphues, K. The *let-99* gene is required for proper spindle orientation during cleavage of the *C. elegans* embryo. *Development* **125,** 1337–1346 (1998).
36. Clark-Maguire, S. & Mains, P. E. *mei-1*, a gene required for meiotic spindle formation in *Caenorhabditis elegans*, is a member of a family of ATPases. *Genetics* **136,** 533–546 (1994).
37. Guo, S. & Kemphues, K. J. A non-muscle myosin required for embryonic polarity in *Caenorhabditis elegans*. *Nature* **382,** 455–458 (1996).
38. Pichler, S. *et al.* OOC-3, a novel putative transmembrane protein required for establishment of cortical domains and spindle orientation in the P_1 blastomere of *C. elegans embryos*. *Development* **127,** 2063–2073 (2000).
39. Guo, S. & Kemphues, K. J. *par-1*, a gene required for establishing polarity in *C. elegans* embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. *Cell* **81,** 611–620 (1995).
40. Hung, T. J. & Kemphues, K. J. PAR-6 is a conserved PDZ domain-containing protein that colocalizes with PAR-3 in *Caenorhabditis elegans* embryos. *Development* **126,** 127–135 (1999).
41. Raich, W. B., Moran, A. N., Rothman, J. H. & Hardin, J. Cytokinesis and midzone microtubule organization in *Caenorhabditis elegans* require the kinesin-like protein ZEN-4. *Mol. Biol. Cell* **9,** 2037–2049 (1998).
42. Powers, J., Bossinger, O., Rose, D., Strome, S. & Saxton, W. A nematode kinesin required for cleavage furrow advancement. *Curr. Biol.* **8,** 1133-1136 (1998).
43. Matthews, L. R., Carter, P., Thierry, M. D. & Kemphues, K. ZYG-9, a *Caenorhabditis elegans* protein required for microtubule organization and function, is a component of meiotic and mitotic spindle poles. *J. Cell Biol.* **141,** 1159-1168 (1998).
44. Carter, P. W., Roos, J. M. & Kemphues, K. J. Molecular analysis of *zyg-11*, a maternal-effect gene required for early embryogenesis of *Caenorhabditis elegans*. *Mol. Gen. Genet.* **221,** 72–80 (1990).

Supplementary information is available on *Nature*'s World-Wide Web site (http://www.nature.com) or as paper copy from the London editorial office of *Nature*.

Acknowledgements

For help in improving the manuscript, we thank A. Desai, A. Ephrussi, S. Grill, M. Labouesse, I. Mattaj and B. Sönnichsen. Work in the Hyman laboratory is supported by the EMBL (European Molecular Biology Laboratory) and the MPI (Max-Planck-Institute). P.G. was supported by a fellowship from the Swiss National Science Foundation. A.C. is supported by the UK Medical Research Council.

Correspondence and requests for materials should be addressed to P.G (e-mail: Pierre.Gonczy@isrec.unil.ch) or A.A.H. (e-mail: hyman@EMBL-Heidelberg.de). The complete data set, including sample time-lapse recordings, has been deposited in an open-access database (http://mpi-web.embl-heidelberg.de/dbScreen/).