Note

Identification of 1088 New Transposon Insertions of *Caenorhabditis elegans*: A Pilot Study Toward Large-Scale Screens

Edwige Martin,* Hélène Laloux,* Gaëlle Couette,* Thierry Alvarez,* Catherine Bessou,* Oliver Hauser,* Satis Sookhareea,† Michel Labouesse† and Laurent Ségalat*,[1]

**CGMC, CNRS-UMR 5534, Université Lyon1, 69100 Villeurbanne, France and †IGBMC, BP. 163, 67404 Illkirch, France*

Manuscript received December 26, 2001
Accepted for publication June 18, 2002

ABSTRACT

We explored the feasibility of a strategy based on transposons to generate identified mutants of most *Caenorhabditis elegans* genes. A total of 1088 random new insertions of *C. elegans* transposons Tc1, Tc3, and Tc5 were identified by anchored PCR, some of which result in a mutant phenotype.

PROJECTS to build large-scale collections of mutants exist for several model organisms amenable to genetics, including yeast, flies, and mouse (Zambrowicz *et al.* 1998; Liao *et al.* 2000; Vidan and Snyder 2001). In *Caenorhabditis elegans*, there are currently mutants for <10% of the estimated 19,000 genes of *C. elegans*. In an approach complementary to knockouts, nondirected transposon-based insertional mutagenesis can be used to generate collections of mutants. In the mutants obtained, insertion sites are determined *a posteriori* by molecular analysis. *C. elegans* Tc elements from the *mariner* family have been used extensively for both forward and reverse genetics purposes (Anderson 1995; Plasterk and van Luenen 1997). However, a systematic study of Tc distribution at the genome level in mutator strains and of their potential mutagenic effect is not available. In this study, we characterized random insertions of Tc1 and Tc3 elements of the mariner superfamily (Plasterk *et al.* 1999) and of the distantly related Tc5 element (Collins and Anderson 1994) to assay the feasibility of such an approach to generate a large-scale collection of identified mutants of *C. elegans*.

A technical difficulty in *C. elegans* is that natural elements used for mutagenic purposes exist in multiple copies in all strains. Approximately 30 copies of Tc1, 20 copies of Tc3, and 6 copies of Tc5 are present in the reference N2 strains. As a consequence, molecular screens need to distinguish between new and preexisting transposon copies in genomes to be analyzed.

Detection of the insertions: To generate new insertions of Tc1, Tc3, and Tc5, we propagated strains carrying the *mut-7* mutation (Ketting *et al.* 1999) over 10 generations. Worms were frozen to be included in the strain collection, and an aliquot was used to extract DNA. We detected insertions by a modification of the transposon display protocol (Wicks *et al.* 2000), such that the radioactive and polyacrylamide gel steps were removed. Worms (15 μl) were collected in M9 buffer, frozen at −80° for at least 30 min, and incubated in 100 mM Tris-HCl (pH 8.5), SDS 1%, EDTA 50 mM, NaCl 0.1 M, and 100 μg/ml proteinase K for 3 hr at 65°. DNA was extracted with phenol/chloroform, precipitated with 0.1 vol of sodium acetate and 0.8 vol of isopropanol, and resuspended in 50 μl TE (10 mM Tris-HCl pH 7.6, 1 mM EDTA) + 10 μg/ml RNAse. Then 5 μl of genomic DNA was digested with restriction enzymes. For enzymes producing cohesive ends, fragments were blunted with the Klenow enzyme. After heat inactivation of the enzymes, a linker formed of 5′-ACCTGCCC-3′ and 5′-CTA ATACGACTCACTATAGGGCTCGAGCGGCCGCCCGG GCAGGT-3′ was ligated to the fragments. A nested PCR was performed on 0.5 μl of the ligation reaction, using primer AP1 and one of the Tc primers (see below) for the first round of PCR (20 cycles). The second round of PCR was performed on 0.01 μl of the first PCR product using primer AP2 and one of the Tc primers (30 cycles). PCRs were performed with the Eurobio enzyme (Evry, France) in the buffer provided by the manufacturer.

Primer sequences are as follows: AP1, CCATCCTAAT ACGACTCACTATAGGGC; AP2, ACTCACTATAGGGC TCGAGCGGC; TC1R1S, GATCGACTCGATGCCACGT CGTTG; TC1R2, GATTTTGTGAACACTGTGGTGAAG; TC3.1, GGTCCTATAGAAGTTTCACACTGG; TC3.2, TTC

[1]*Corresponding author:* CGMC, Université Lyon1, 43 bld du 11 Novembre, 69622 Villeurbanne Cedex, France.
E-mail: segalat@maccgmc.univ-lyon1.fr

GGAAGTTCCTCAAACCTTC; TC5.1, GCCAAACCTGCT CTGAAGCAG; and TC5.4, GGATCATCTGTAACTATCC TCTATCG.

Annealing temperature was 58° for Tc1 and 48° for Tc3 and Tc5. PCR products were run on 2% agarose gels. Bands that were unique to a clone were picked with a plastic inoculator, quickly vortexed in 100 μl water, and re-PCRed using the set of primers used for the second round of PCR. The PCR products were purified on Promega (Madison, WI) Wizard columns and sequenced.

Each DNA sample was analyzed independently with three enzymes. Most sequences reported in this study were obtained using enzymes *Cla*I, *Hin*dIII, and *Pml*I. Only a fraction of the insertions were detected by each enzyme. On the basis of the number of insertions detected with each enzyme, we estimate that ~2/3 of the insertions are detected using this protocol.

A total of 862 clones were generated. On average, two extra bands per clone (corresponding to new insertions) were observed, reamplified, and sequenced. Sequences were subsequently compared to the *C. elegans* genome by Blast. A total of 1088 sequences resulted in unambigous Blast results (625 Tc1, 253 Tc3, 210 Tc5). The list of insertions can be found on http://cgmc.univ-lyon1.fr and on http://www.wormbase.org.

To validate the protocol, we assayed whether we could recover predicted insertions from frozen worms. This is a critical point because (1) the anchored-PCR technique might produce artifacts and (2) original plates might carry a mixture of heterozygous and homozygous animals. We designed insertion-specific primers located ~500 bp from the predicted insertion sites. Thawed worms were singled out and PCR tested using as primers the Tc-specific primer that had been used in the first instance and the insertion-specific primer. In this way, we could recover an insertion in 64 of 69 strains tested. In most cases, more than half of the worms carried the insertion.

Distribution of the insertions: One goal of this study was to observe the distribution of randomly generated Tc insertions in the genome. The distribution of each transposon (Tc1, Tc3, and Tc5) is shown on the physical map (Figure 1). No gross bias for any of the three transposons is observed, and we detected insertions in all chromosomic regions. The greatest disequilibrium between chromosomes is seen for Tc1, for which chromosome V has three times as many insertions as chromosome III (two times if standardized to chromosome length). The current resolution is not sufficient to determine if some genes are more prone to transposon integration than others.

Another important parameter of transposon distribution is the type of genomic region in which the transposons jump (exons, introns, intergenic regions). We observed that ~20% of the insertions are located in coding sequences, 30% in introns, and 5% in 5′- or 3′-untranslated regions (arbitrarily defined as 100 bp before and after the coding sequence as defined in ACeDB/Wormbase databases). The remainder are located in intergenic regions. Coding sequence accounts for 26% of the *C. elegans* genome and introns for 14% (*C. elegans* Consortium 1998). Results were similar for each of the transposons. A likely explanation for the high proportion of intronic insertions comes from the AT-rich content of introns (the target sequence of Tc1 and Tc3 is TA, whereas Tc5 recognizes TNA).

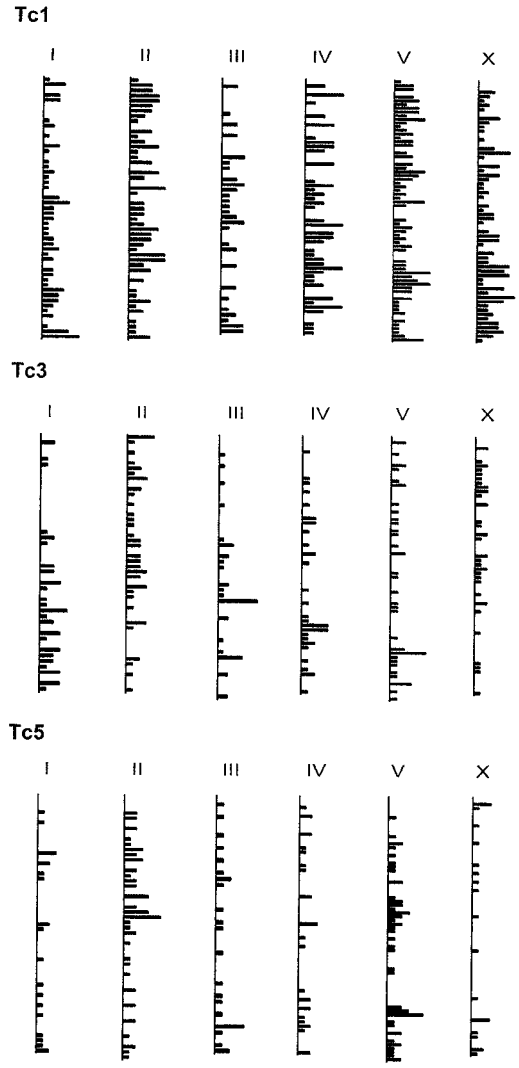

Figure 1.—Distribution of insertions on the *C. elegans* physical map. Tc1, Tc3, and Tc5 are shown separately. Each chromosome is shown in the standard orientation. The length of each bar indicates the number of insertions within a 300-kb interval.

Mutagenic properties of insertions located in exons: We next tried to assay the mutagenic properties of insertions located in exons by asking whether these insertions produced a phenotype comparable to that resulting from a reduction of gene function. The loss-of-function phenotype of 6 out of the 277 exon insertions was known, as a result of either genetic (4 genes) or RNAi (2 genes) stud-

TABLE 1

Phenotype associated with homozygous insertion alleles

Allele name	Tc type	Gene name	Mutant name	Phenotype mutant or RNAi[a]	Phenotype insertion
cxP3316	Tc5	K12F2.1	*myo-3*	95% dead eggs. Escapers paralyzed (1, 2)	50% dead eggs. Escapers paralyzed
cxP3318	Tc3	F42A6.8	—	Constipated (1)	Constipated
cxP3915	Tc1	F10G8.4	—	20% dead eggs (1)	No visible phenotype
cxP4306	Tc1	R13A5.3	—	Hyperactive (1)	Hyperactive
cxP4317	Tc1	Y16B4A.1	*unc-3*	Unc (1, 3)	No visible phenotype
cxP1016	Tc1	F33H2.5	—	100% dead eggs (1, 4)	No visible phenotype
cxP3027	Tc1	F53C3.7	—	50% dead eggs (1)	25% dead eggs
cxP4238	Tc1	C17C3.11	—	50% sterile (1)	25% dead eggs
cxP5915	Tc1	F59B10.1	—	Sluggish and unhealthy (1, 5)	Sluggish
cxP8142	Tc3	M03D4.6	—	Sterility (1)	Larval arrest
cxP6814	Tc1	H24O09.2	—	Sterility (1)	No visible phenotype
cxP8133	Tc3	C23H3.4	—	Larval arrest (1)	No visible phenotype
cxP4925	Tc5	F14F3.1	*vab-3*	Notched head (6)	No visible phenotype
cxP5102	Tc1	B0240.3	*daf-11*	daf-c (7)	20% dauers

[a] References: (1) This study; (2) Waterston (1989); (3) Prasad *et al.* (1998); (4) Fraser *et al.* (2000); (5) Maeda *et al.* (2001); (6) Lewis and Hodgkin (1977); (7) Golden and Riddle (1984).

ies. We performed RNAi on 92 additional genes that were chosen at random. Ten genes produced obvious phenotypes by RNAi (Table 1), a ratio similar to that published in systematic screens (Fraser *et al.* 2000; Gonczy *et al.* 2000; Piano *et al.* 2000; Hanazawa *et al.* 2001; Maeda *et al.* 2001). The corresponding 16 clones were thawed, and the insertion could be recovered in 14 of them. Homozygous insertions led to phenotypes resembling those obtained by mutation or inactivation of the gene in 7 out of 14 cases (Table 1). In most cases, the phenotype of the insertion was weaker than the one observed by mutation or RNAi, a feature that is common to transposon-induced mutations (Anderson 1995).

The mutagenic properties of randomly generated transposon insertions in exons in *C. elegans* have been a matter for discussion for many years. It has been clearly demonstrated that insertions located in the coding sequence of several genes do not lead to a phenotype because the transposon is spliced out of the mRNA (Rushforth and Anderson 1996), but it is unclear how general this phenomenon is. Although *n* is limited, our results indicate that Tc insertions in exons are not silent; half of randomly generated insertions located in exons of genes having a visible phenotype led to a similar phenotype.

This work was designed as a pilot experiment to analyze the feasibility of a large-scale production of mutants based on *C. elegans* Tc elements. It performs a community service by providing novel Tc insertions in or near 600 genes. In addition, the issues addressed in this study will be relevant to other types of transposons that may be used as mutagenic tools in *C. elegans* in the future.

The authors are grateful to J. Godet, J. Samarut, J.-A. Lepesant, M. Philippe, and P. Couble for their continuous support for this project. We also thank T. Drynda for help with the web site, S. Wicks for the transposon display protocol, and L. Stein for incorporating the data into public databases. This work was supported by the Centre National de la Recherche Scientifique (CNRS), the Rhône-Alpes district, and the Association Française pour la Recherche contre le Cancer (ARC).

LITERATURE CITED

Anderson, P., 1995 Mutagenesis, pp. 31–58 in *Caenorhabditis elegans*, edited by H. F. Epstein and D. C. Shakes. Academic Press, San Diego.

C. elegans Consortium, 1998 Genome sequence of the nematode *C. elegans*: a platform for investigating biology. Science **282:** 2012–2017.

Collins, J., and P. Anderson, 1994 The Tc5 family of transposable elements in *Caenorhabditis elegans*. Genetics **137:** 771–781.

Fraser, A. G., R. S. Kamath, P. Zipperlen, M. Martinez-Campos, M. Sohrmann *et al.*, 2000 Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference. Nature **408:** 325–330.

Golden, J. W., and D. L. Riddle, 1984 A pheromone-induced developmental switch in *Caenorhabditis elegans*: temperature-sensitive mutants reveal a wild-type temperature-dependent process. Proc. Natl. Acad. Sci. USA **81:** 819–823.

Gonczy, P., G. Echeverri, K. Oegema, A. Coulson, S. Jones *et al.*, 2000 Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. Nature **408:** 331–336.

Hanazawa, M., M. Mochii, N. Ueno, Y. Kohara and Y. Iino, 2001 Use of cDNA subtraction and RNA interference screens in combination reveals genes required for germ-line development in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA **98:** 8686–8691.

Ketting, R., T. Haverkamp, H. van Luenen and R. Plasterk, 1999 *Mut-7* of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. Cell **99:** 133–141.

Lewis, J. A., and J. A. Hodgkin, 1977 Specific neuroanatomical changes in chemosensory mutants of the nematode *Caenorhabditis elegans*. J. Comp. Neurol. **172:** 489–510.

Liao, G., E. Rehm and G. Rubin, 2000 Insertion site preferences of the P transposable element in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **97:** 3347–3351.

Maeda, I., Y. Kohara, M. Yamamoto and A. Sugimoto, 2001 Large-scale analysis of gene function in *Caenorhabditis elegans* by high-throughput RNAi. Curr. Biol. **11:** 171–176.

Piano, F., A. J. Schetter, M. Mangone, L. Stein and K. J. Kemphues, 2000 RNAi analysis of genes expressed in the ovary of *Caenorhabditis elegans*. Curr. Biol. **10:** 1619–1622.

Plasterk, R., and H. van Luenen, 1997 Transposons, pp. 97–116 in *C. elegans II*, edited by D. Riddle, T. Blumenthal, B. Meyer and J. Priess. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Plasterk, R., Z. Izsvak and Z. Ivics, 1999 Resident aliens. Trends Genet. **15:** 326–332.

Prasad, B. C., B. Ye, R. Zackhary, K. Schrader, G. Seydoux *et al.*, 1998 *unc-3*, a gene required for axonal guidance in *Caenorhabditis elegans*, encodes a member of the O/E family of transcription factors. Development **125:** 1561–1568.

Rushforth, A., and P. Anderson, 1996 Splicing removes the *Caenorhabditis elegans* transposon Tc1 from most mutant pre-mRNAs. Mol. Cell. Biol. **16:** 422–429.

Vidan, S., and M. Snyder, 2001 Large-scale mutagenesis: yeast genetics in the genome era. Curr. Opin. Biotechnol. **12:** 28–34.

Waterston, R. H., 1989 The minor myosin heavy chain, mhcA, of *Caenorhabditis elegans* is necessary for the initiation of thick filament assembly. EMBO J. **8:** 3429–3436.

Wicks, S., C. de Vries, H. van Luenen and R. Plasterk, 2000 CHE-3, a cytosolic dynein heavy chain, is required for sensory cilia structure and function in *Caenorhabditis elegans*. Dev. Biol. **221:** 295–307.

Zambrowicz, B., G. Friedrich, E. Buxton, S. Lilleberg, C. Person *et al.*, 1998 Disruption and sequence identification of 2,000 genes in mouse embryonic stem cells. Nature **392:** 608–611.

Communicating editor: P. Anderson