The *unc-22(IV)* Region of *Caenorhabditis elegans*: Genetic Analysis of Lethal Mutations

Denise V. Clark, Teresa M. Rogalski,[1] Leo M. Donati and David L. Baillie

Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6

Manuscript received November 10, 1987
Accepted February 11, 1988

ABSTRACT

The organization of essential genes in the *unc-22* region, defined by the deficiency *sDf2* on linkage group IV, has been studied. Using the balancer *nT1 (IV;V)*, which suppresses recombination over 49 map units, 294 lethal mutations on LGIV(right) and LGV(left) were recovered using EMS mutagenesis. Twenty-six of these mutations fell into the *unc-22* region. Together with previously isolated lethal mutations, there is now a total of 63 lethal mutations which fall into 31 complementation groups. Mutations were positioned on the map using eight overlapping deficiencies in addition to *sDf2*. The lethal alleles and deficiencies in the *unc-22* region were characterized with respect to their terminal phenotypes. Mapping of these lethal mutations shows that *sDf2* deletes a minimum of 1.8 map units and a maximum of 2.5 map units. A minimum estimate of essential gene number for the region using a truncated Poisson calculation is 48. The data indicate a minimum estimate of approximately 3500 essential genes in the *Caenorhabditis elegans* genome.

OUR laboratory is interested in revealing how genes essential for development are organized at both the genetic and molecular levels in the genome of the nematode *Caenorhabditis elegans*. Several regions of the *C. elegans* genome have been genetically characterized with respect to essential genes in the past, particularly on linkage group (LG) I (ROSE and BAILLIE 1980; HOWELL *et al.* 1987), LGII (SIGURDSON, SPANIER and HERMAN 1984), LGV (ROSENBLUTH *et al.* 1988), LGX (MENEELY and HERMAN 1979, 1981) and LGIV around *ama-1* (ROGALSKI and RIDDLE 1988) as well as in the *unc-22* gene cluster (ROGALSKI, MOERMAN and BAILLIE 1982; ROGALSKI and BAILLIE 1985) which is the subject of this study.

The *unc-22* gene codes for a muscle component. It has been characterized genetically (MOERMAN and BAILLIE 1979; ROGALSKI and BAILLIE 1985) and has been cloned (MOERMAN, BENIAN and WATERSTON 1986). In addition to the cloning of *unc-22*, restriction fragment length differences (RFLDs) have been genetically mapped and cloned in the 2.5-map-unit interval surrounding *unc-22*, between *unc-43* and *unc-31* (BAILLIE, BECKENBACH and ROSE 1985). With the continued addition of overlapping cosmids to the original clones (COULSON *et al.* 1986), there is now a total of about 1,100 kilobases of cloned DNA in the *unc-43* to *unc-31* interval. However, the cosmids are not yet contiguous throughout the region.

In the previous work on the *unc-22* region (ROGALSKI, MOERMAN and BAILLIE 1982; ROGALSKI and BAILLIE 1985), 20 essential genes were identified in the 2 map unit interval defined by the deficiency *sDf2*. From these studies, it was estimated that there are 32 essential genes in the *unc-22* region, based on the Poisson distribution of the lethal mutations. Therefore, the region had not been saturated with mutations in all essential genes. The reasons for a continued genetic characterization of the *unc-22* region are threefold. First, approaching saturation of the essential genes with lethal mutations allows one to make a more accurate estimate of gene number in the *unc-22* region. Second, identification of all essential genes, along with the positioning of deficiency breakpoints on the genetic map, is valuable for molecular characterization of the region. Third, a collection of alleles of each essential gene provides information about its pattern of expression in development and perhaps about the relationships between adjacent genes in the region.

In this paper we describe the isolation of a set of EMS induced lethal mutations using the balancer *nT1(IV;V)* (FERGUSON and HORVITZ 1985). Along with an updated map of essential genes in the *sDf2* region and mapping data for newly identified essential genes, we present a compiled list of all lethal mutations in the *sDf2* region isolated to date, along with their terminal phenotypes. By using the distribution of allele number for the essential genes in the *sDf2* region, we arrive at a Poisson estimate of total essential gene number for this region. The extent of the *nT1* balanced region is refined to provide an estimate of essential gene number in the balanced region. This estimate is derived from determining the pro-

[1] Present address: Division of Biological Sciences, University of Missouri, Columbia, Missouri 65211.

portion of lethal mutations (isolated in the present screen) that fall into the *sDf2* region and from the Poisson estimate of essential gene number in the *sDf2* region.

MATERIALS AND METHODS

The genetic nomenclature in this paper follows the recommendations of HORVITZ *et al.* (1979).

Stock maintenance and strains: Nematodes were maintained on petri plates containing nematode growth medium (NGM) streaked with *Escherichia coli* OP50 (see BRENNER 1974).

The wild-type strain N2 (var. Bristol) and strains carrying the following mutations used in this study were obtained from the stock collection at the Medical Research Council, Cambridge, England, or from the Caenorhabditis Genetics Center at the University of Missouri, Columbia. LGIV: *dpy-4(e1166)*, *dpy-9(e424)*, *dpy-13(e184)*, *dpy-20(e1282)*, *mDf7* (ROGALSKI and RIDDLE 1988), *mec-3(e1338)*, *unc-5(e152)*, *unc-17(e933)*, *unc-31(e169)*, *unc-33(e204)*; LGV: *dpy-11(e224)*, *dpy-21(e428)*, *unc-76(e911)*; *nT1(IV;V)* (FERGUSON and HORVITZ 1985): from the strain CB3608 which has the genotype *dpy-20(e1282) dpy-26(n199)/nT1(IV)*; +/*nT1(V)*. The strains carrying *nDf27(IV)* and *lin-1(e1777) IV* and the strain *lin-3(n378) unc-22(e66)* were obtained from the R. Horvitz laboratory, M.I.T. The remaining LGIV mutations were isolated at Simon Fraser University: *unc-22(s7)* (MOERMAN and BAILLIE 1979), *sDf2*, *sDf7*, *sDf8*, *sDf9*, *sDf10* (MOERMAN and BAILLIE 1981), *sDf21* and *sDf22* (G. WILD and D. BAILLIE, unpublished results). The *sDf2* region lethal mutations on LGIV isolated previously at Simon Fraser University and those isolated in the screen described below are listed in Table 2 with corresponding references.

It should be noted that throughout this work the *unc-22(s7)* mutation was used as a marker (MOERMAN and BAILLIE 1979). It is a conditionally semidominant mutation where the *unc-22(s7)*/+ individual has a "twitcher" phenotype in a 1% nicotine solution (Sigma) and is otherwise phenotypically wild type. Individuals that are homozygous wild type are contracted and paralyzed in nicotine, whereas *unc-22(s7)* homozygotes are unconditional twitchers.

Screen for recessive lethal mutations: Young gravid hermaphrodites of the genotype *unc-22 unc-31/nT1(IV)*; +/*nT1(V)* were exposed to 0.012 M EMS for 4 hr. P_0 hermaphrodites were then allowed to lay eggs for two 24-hr broods. This procedure prevented overcrowding and thus ensured that all F_1s were recovered from each parent over the 48-hr period. All wild type F_1s were then individually allowed to self-fertilize and the F_2s were screened for the absence of gravid Unc-22 Unc-31 individuals. At least three wild-type individuals from each strain carrying a putative recessive lethal mutation were selected. The progeny of these were screened for the absence of Unc-22 Unc-31 recombinants to confirm that the mutation lies within the *nT1* balanced region.

To determine the induction frequency of spontaneous recessive lethal mutations over *nT1*, the same protocol was used as described above, omitting the 0.012 M EMS.

Extent of the *nT1* balanced region on LGIV and LGV: In order to determine the extent of recombination suppression by *nT1* on LGIV and LGV more precisely, various pairs of morphological mutations were placed in *trans* to *nT1* and to a wild-type N2 chromosome for comparison. Table 1 shows the various double mutant chromosomes used. These were constructed as described by BRENNER (1974). Double mutants, heterozygous with *nT1*, were constructed in the following manner. Homozygous double mutant (*a b*) hermaphrodites were mated to *unc-22 unc-31/nT1(IV)*; +/*nT1(V)* males. Double mutant heterozygotes [*a b*/*nT1(IV)*; +/*nT1(V)* or +/*nT1(IV)*;a *b*/*nT1(V)*] were selected in a 1% nicotine solution as wild type individuals lacking the twitching phenotype. The pseudolinkage between LGIV and LGV caused by *nT1* allowed construction of these strains.

Mapping and complementation tests: Recombination mapping was carried out at a constant 20 degrees after the recommendations of ROSE and BAILLIE (1979). Initial mapping of all lethal mutations isolated in the screen to either LGIV or LGV involved crossing each strain to N2 males and scoring the progeny of one or two hermaphrodites of the genotype (*let-x*) *unc-22 unc-31*/+ + +(*IV*); +/+(*V*) or *unc-22 unc-31*/+ +(*IV*);*let-x*/+(*V*). A greater than 3:1 ratio of wild type to Unc-22 Unc-31 progeny indicated linkage to LGIV. All LGIV lethal mutations were retained and those within 4 map units from *unc-22* were positioned further through complementation tests with *mDf7*, *sDf2* and *sDf21*. Positioning of lethal mutations within *sDf2* was accomplished through complementation tests with a set of deficiencies whose breakpoints divide the *sDf2* region into ten zones (see Figure 2). Once a mutation was positioned into a zone, it was assigned to a complementation group through complementation tests with alleles of all genes within that zone. For complementation tests with alleles of previously identified genes that had multiple alleles (see Table 2), only one representative allele was used (see RESULTS for further discussion). The positions of recently identified essential genes were confirmed and refined by three factor mapping. The progeny of several (*let-x*) *unc-22 unc-31*/+ + + individuals were scored (see Table 3).

Effective lethal stages: Effective lethal stages of homozygotes and hemizygotes for EMS induced lethal mutations and of homozygotes for deficiencies in the *sDf2* region were determined. For homozygous lethal phenotypes arresting later than the early larval stage, heterozygous parents from the stock strain were selfed at 20°, three Unc-22 Unc-31 larvae or sterile adults were selected and their average length was measured 3 days later (to ensure completion of growth) under a dissecting microscope (accuracy ± 0.05 mm). For lethal mutations arresting in the early larval stage, putative Unc-22 Unc-31 homozygotes could not be easily identified using the *nT1* balanced strain due to the occurrence of egg or early larval stage arrested worms that are probably *nT1* segregation aneuploids. Therefore, for these mutations, (*let-x*) *unc-22 unc-31/nT1(IV)*; +/nT1(V) or *Df/nT1* hermaphrodites were mated to *unc-22 unc-31/nT1(IV)*; +/*nT1(V)* males, gravid Unc-22 Unc-31 F_1 progeny were selected and allowed to lay eggs in 5–12-hr broods. As viable progeny matured, they were removed from the plate until only developmentally arrested individuals were left. These were measured if they had hatched from the egg. To determine the hemizygous phenotype of EMS induced lethal mutations over *sDf2*, (*let-x*) *unc-22 unc-31/sDf2* individuals were selected and measured. If the stage of arrest was in the egg or early larva, it could not be determined since there could be some developmentally arrested aneuploids among the parental hermaphrodite [*sDf2/nT1(IV)*; +/*nT1(V)*] self progeny.

The larval stage of arrest was determined by comparing the length of mutant individuals to the *unc-22(s7)* growth curve (ROGALSKI, MOERMAN and BAILLIE 1982). It should be noted that the developmental stage of arrest for larval lethal mutations was not classified as one of L1, L2, L3 or L4 because no attempt was made to observe and record

TABLE 1

***nT1* recombination suppression**

Linkage group	Interval a	Interval b	*a b/+ +*[a] A	B	Total	Distance	*a b/nT1*[b] A	B	Total	Distance
IV	*dpy-9*	*unc-17*	247	197	1715	31 (29–33)	71	43	462	35 (30–40)
	unc-17	*dpy-4*	152	159	2244	15 (14–16)	2	0	3568	~0
	unc-22	*dpy-4*				5[c]	1	0	1172	~0
	unc-33	*dpy-13*	20	27	1892	2.5 (1.8–3.2)	0	0	1073	0
	lin-1	*dpy-13*	47	67	1766	7.8 (6.0–9.6)[d]	1	0	724	~0
V	*dpy-11*	*unc-76*	75	48	1784	8.6 (6.8–10.4)[d]	0	0	1243	0
	unc-76	*dpy-21*	60	64	1069	12 (10–14)	25	44	701	11 (9–13)

[a] Distance in map units calculated from $100p$ where $p = 1 - (1 - 2R)^{1/2}$ and R = (A + B)/total progeny scored (BRENNER 1974). The 95% confidence intervals are shown in parentheses.

[b] Distance calculated assuming *nT1* behaves like the reciprocal translocation *eT1(III; V)* (*see* ROSENBLUTH and BAILLIE 1981) that segregates inviable aneuploids so that $p = 0.6–0.6[1 - (10/3)R]^{1/2}$ and R = (A + B)/total progeny scored.

[c] From EDGLEY and RIDDLE 1987.

[d] Dpy recombinants only used to calculate $p = 1 - (1 - 3R')^{1/2}$ and R' = A/(A + number of wild types).

the larval moults. Instead, we classified the length of an individual as early, mid or late larval based on its position on the growth curve. The growth of *unc-22(s7) unc-31(e169)* individuals follows the same curve as for *unc-22(s7)* individuals at 20° (data not shown).

RESULTS

Screen for lethal mutations: Using the screening protocol described in MATERIALS AND METHODS, we established a recessive lethal mutation frequency for the *nT1* balanced region. Out of 3398 F_1 "chromosomes" [LGIV(right) and LGV(left)] screened, 294 lethal mutations were recovered in the *nT1* balanced region (see Figure 1), giving a recessive lethal mutation frequency of 8.7% with 0.012 M EMS.

The types of mutations recovered include egg, larval and adult sterile recessive lethal mutations. Owing to the transparency of Unc-22 Unc-31 individuals and the visibility of fertilized eggs through the body wall, adult sterile mutations were recovered since fertilized eggs were not visible or were greatly reduced in number in individuals carrying these mutations. Any maternal effect mutations leading to fertilized eggs that subsequently fail to develop would thus not be recovered in this screen.

The relative lethal mutation frequencies on LGIV(right) and LGV(left) were 4.9% (169 mutations) and 3.8% (125 mutations) respectively. There is a probability that some of the strains carrying lethal mutations on LGIV also carry a mutation on LGV, since this chromosome was not marked with a morphological mutation. The probability of this occurring is low (approximately 6 out of the 169 LGIV mutations should carry an additional lethal mutation on LGV).

In order to establish the number of lethal mutations that are EMS induced *vs.* those that are spontaneous, a screen was conducted in the absence of EMS. Using the same strain as was used in the EMS screen, *unc-22 unc-31/nT1(IV); +/nT1(V)*, two spontaneous lethal mutations were isolated out of 1736 chromosomes screened. One mutation mapped 17 map units from *unc-22* on LGIV and the other mapped in the region deleted by *sDf21* but not by *sDf2* to the right of *unc-22*. The spontaneous lethal mutation frequency in the *nT1* balanced region is about 0.1%. Therefore, out of the 294 mutations isolated in the EMS screen, about four lethal mutations are expected to be of spontaneous origin.

Extent of the *nT1* balanced region: Originally, FERGUSON and HORVITZ (1985) determined that *nT1* caused pseudolinkage between LGIV(right) and LGV(left) and suppressed recombination in the *unc-17* to *dpy-4* and *unc-60* to *dpy-11* intervals (see Figure 1). The pseudolinkage between LGIV(right) and LGV(left) was observed in *nT1* heterozygote self progeny. All aneuploid zygotes resulting from various combinations of half translocations and normal chromosomes appeared to arrest development in the egg. The euploid progeny consist of homozygotes for normal LGIV and LGV, parental heterozygotes and *nT1* homozygotes (*nT1* is associated with a vulvaless phenotype) in a 1:4:1 ratio, respectively. In order to further refine the extent of *nT1* recombination suppression, other intervals on LGIV and LGV were examined and the results are listed in Table 1. On LGIV, it now appears that *nT1* effectively balances the 21 map unit interval from *lin-1* to *dpy-4* (mapping data includes those from EDGLEY and RIDDLE 1987). It should be noted that there is an occasional Unc recombinant from the *unc-22 dpy-4/nT1(IV); +/nT1(V)* and *unc-17 dpy-4/nT1(IV); +/nT1(V)* individuals. In these cases, the recombinant *unc* chromosome did not carry *nT1*. This was demonstrated by the inability of these recombinant chromosomes to sup-

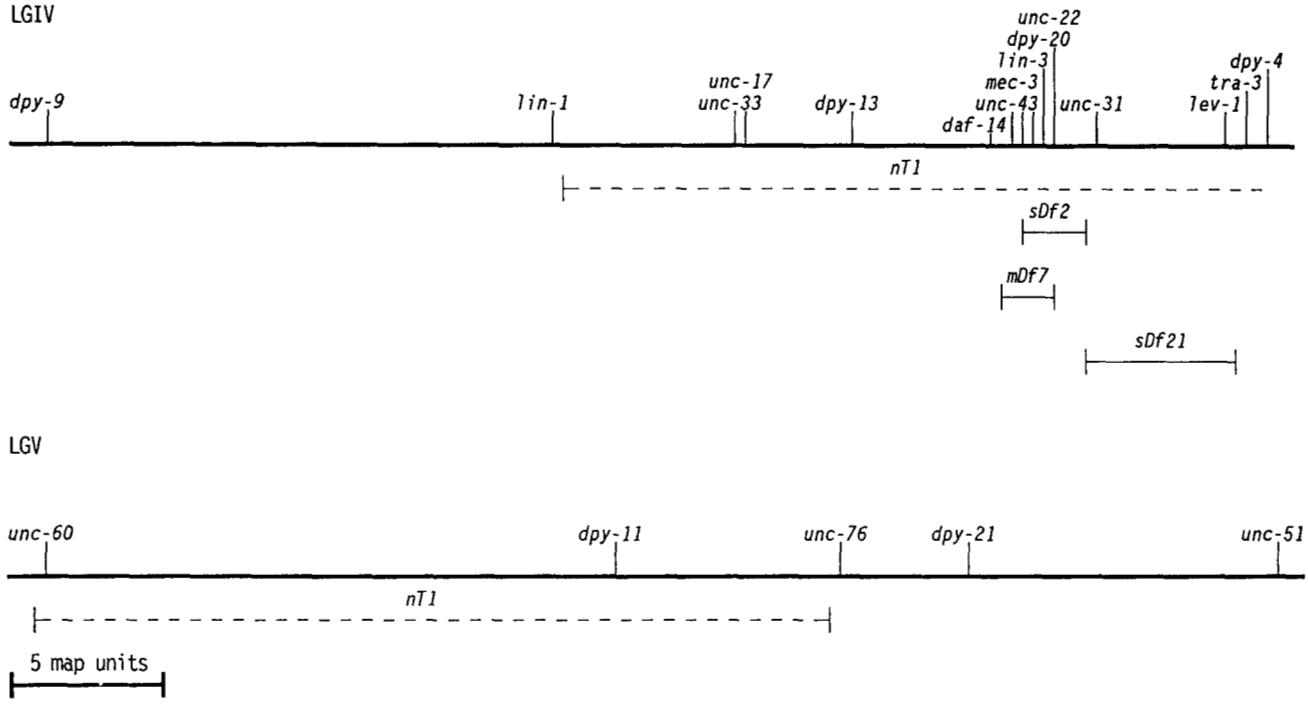

FIGURE 1.—Genetic map of LGIV and LGV showing the region balanced by *nT1(IV;V)*. Nonessential genes of importance to this study are shown with map positions derived from EDGLEY and RIDDLE (1987) and from this study. The overlapping deficiencies *mDf7*, *sDf2* and *sDf21* are shown.

press recombination (data not shown) and by the lack of vulvaless progeny from the Unc recombinants. Therefore, it seems that recombination can occur between *nT1* and the normal LGIV homologue over a small interval to the left of *dpy-4*. This probably means that *nT1* is a more complex rearrangement than a simple reciprocal translocation. It should also be noted that *nT1* is not completely stable in balancing *tra-3*, which lies approximately one map unit to the left of *dpy-4* (J. HODGKIN, personal communication). On LGV, *nT1* suppresses recombination from at least *unc-60* to *unc-76*, an interval of about 28 map units (R. ROSENBLUTH and R. JOHNSEN, personal communication). Therefore, the extent of the *nT1* balanced region is about 49 map units.

Mapping and complementation tests: The 169 strains carrying mutations on LGIV were subjected to further analysis. An approximate two-factor map distance from *unc-22* was determined for each of these lethal mutations (data not shown). Complementation tests were then carried out between the mutations up to four map units from *unc-22* (125 mutations) and the three overlapping deficiencies *mDf7*, *sDf2* and *sDf21*. All 26 lethal mutations failing to complement *sDf2* were 2 map units or less from *unc-22*. None of the mutations within the 2–4-map-unit intervals on either side of *unc-22* were found to lie within *sDf2*. This indicates that the initial approximate mapping of mutations served to decrease the number of necessary complementation tests between lethal mutations and *sDf2*. We could safely assume that, from the sample of 169 mutations, all mutations in the *sDf2* region were identified.

A comparison of the density of lethal mutations recovered in the "*nT1* screen" between *daf-14* and *lev-1* shows that lethal mutations are clustered to the left of *unc-22* relative to the region to the right (data not shown). The clustered region is most likely due to a relatively low frequency of recombination (see DISCUSSION). Until further analysis of lethals on the rest of LGIV(right) is complete, we cannot say whether there are other clustered regions. Alleles of essential genes near *ama-1(IV)* have also been characterized and do not form as dense a cluster as those near *unc-22* (ROGALSKI and RIDDLE 1988).

In order to establish complementation groups, the 26 lethal mutations within *sDf2* were mapped to smaller intervals or zones using the deficiencies *sDf7*, *sDf8*, *sDf9*, *sDf10*, *sDf21*, *sDf22*, *mDf7* and *nDf27*. The *unc-22* region has now been subdivided into 10 zones defined by deficiency breakpoints (Figure 2). To define complementation groups within a zone, representative alleles of previously characterized essential genes were used. The representative allele, in the cases where there were multiple alleles with different phenotypes, was the one which exhibited the earliest developmental arrest or most severe lethal phenotype and would most likely be the null allele. The *inter se* complementation tests (data not shown) show that 11 more essential genes (in addition to *lin-3*; see below) have been identified in the *sDf2* region since ROGALSKI and BAILLIE 1985. Table 2 lists each essential

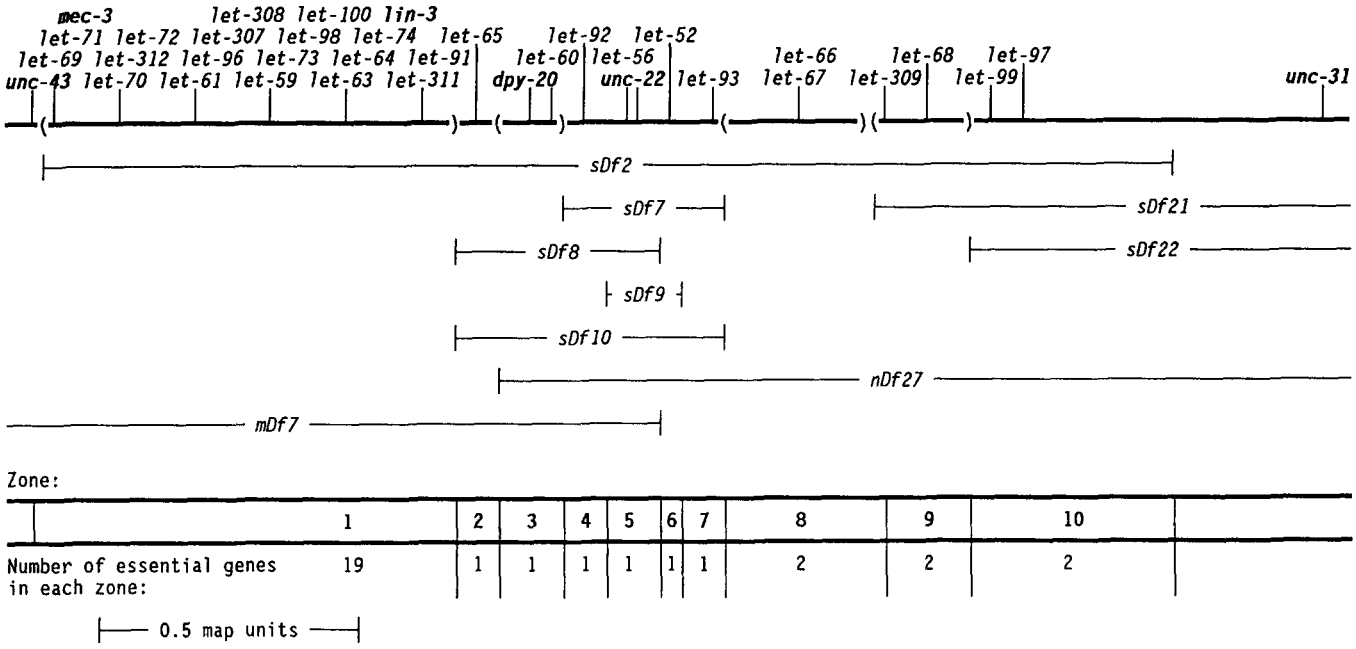

FIGURE 2.—Genetic map of essential genes in the *unc-22* region defined by the deficiency *sDf2*. The 31 essential genes are shown along with the positions of deficiency breakpoints which divide the region into 10 zones. The number of essential genes within each zone is indicated below the map. Brackets on the map indicate that the genes within them have not been positioned relative to each other. The order of genes within the brackets is based on two-factor mapping distances from *unc-22*.

gene, lethal allele, zone, and effective lethal stage for homozygotes and most hemizygotes (*let-x/sDf2*). Some changes in the data from previous reports are as follows. In ROGALSKI and BAILLIE (1985), *s175* was assigned as the only allele of *let-62*, which possibly formed part of a complex locus with *let-59*. Reexamination of these results showed that *s175* and the other *let-59* alleles fail to complement each other. Thus, the gene designation *let-62* has since been discontinued. Also, *s172* was shown to be an allele of *let-69* and not *let-59*, as indicated in ROGALSKI, MOERMAN and BAILLIE (1982).

lin-3 is an essential gene that also has viable alleles with a vulvaless phenotype (FERGUSON and HORVITZ 1985). In the course of this work, *s751* and *s1263* were originally assigned to *let-94*. Subsequently, complementation tests with *lin-3(n378)* (FERGUSON and HORVITZ 1985) showed that *let-94* was also *lin-3*. Therefore, we have discontinued the *let-94* designation and adopted the name *lin-3* for this complementation group.

The mapping data for the eleven new essential genes and *lin-3* are listed in Table 3. Mapping data for all other essential genes can be found in the corresponding references listed in Table 2. The mutations *let-91(s753)*, *let-93(s734)* and *lin-3(s751)* were isolated in a preliminary study using a screen identical to the one described in this paper (DONATI 1985). The remaining ten *let* genes were identified in this study.

The two-factor mapping data of lethal mutations within *sDf2* show that it deletes a region a minimum of 1.8 map units and a maximum of 2.5 map units in size (Figure 2). The left breakpoint of *sDf2* is well defined. *unc-43* has been positioned just outside of *sDf2*, 1.2 map units from *unc-22*, whereas the furthest essential gene to the left of *unc-22* but still within *sDf2* is *let-69*, 1.2 map units from *unc-22* (ROGALSKI and BAILLIE 1985). The right breakpoint of *sDf2* is less clearly defined. This is perhaps due to the relative sparsity of genes to the right of *unc-22*. *let-97*, 0.6 map unit to the right of *unc-22*, is the furthest essential gene from *unc-22* that lies within *sDf2*, while *unc-31* lies 1.3 map units from *unc-22* and is the most proximal gene outside of the right *sDf2* breakpoint. Therefore, the right *sDf2* breakpoint is between 0.6 and 1.3 map units from *unc-22*.

There is one locus that shows a complex complementation pattern. There are two basic explanations for this phenomenon: either some of the alleles are small deletions spanning two complementation groups or some pairs of alleles exhibit intragenic complementation. This locus, *let-91*, has been considered a single, complex gene for the purpose of Poisson analysis of the data. There are four alleles of *let-91*, two of which arrest in mid-larval stages when homozygous and which complement each other (*s678* and *s1185*). The allele *s753*, on the other hand, arrests late in larval development and fails to complement both *s678* and *s1185*. Because *s753* arrests later in development than the other two alleles, it seems unlikely that it is a deletion of two distinct,

TABLE 2

Lethal alleles in the *unc-22* region

Zone[a]	Gene	Allele	Reference[b]	Phenotype of lethal arrest[c] *let-x/let-x*	Phenotype of lethal arrest[c] *let-x/sDf2*
6	*let-52*	*s42*	1	Early larva	Early larva
5	*let-56*	*s46*	1	Late larva	Mid larva
		s50, s168	1	Late larva	ND
		s173	1	Mid larva	Early larva
		s1192, s1210, s1223, s1262		Mid larva	Mid larva
1	*let-59*	*s49*	1	Early larva	Early larva
		s175	1	Multiple stages	Egg to early larva[d]
		s681	2	Early larva	ND
		s1174, s1197		Early larva	Egg to early larva
3	*let-60*	*s59*	1	Mid larva (leaky)	Mid larva
		s1124		Early larva	Egg to early larva
		s1155		Late larva	Mid larva
1	*let-61*	*s65*	1	Late larva	Late larva
1	*let-63*	*s170*	1	Mid larva	Mid larva
		s679	2	Late larva	ND
1	*let-64*	*s171*	1	Sterile (leaky)	Sterile
		s216	1	Sterile	Sterile
2	*let-65*	*s174, s254*	1	Mid larva	Mid larva
		s694	2	Sterile	ND
		s1154, s1222		Mid larva	Mid larva
8	*let-66*	*s176*	1	Early larva	Mid larva
8	*let-67*	*s214*	1	Sterile	Mid larva
9	*let-68*	*s680*	2	Sterile	Mid larva
		s693	2	Late larva	ND
		s696	2	Sterile	ND
		s698	2	Mid larva	ND
		s1258		Early larva	Early larva
1	*let-69*	*s172*	1	Early larva	ND
		s684	2	Late larva	Mid larva
		s1111		Early larva	Egg to early larva
1	*let-70*	*s689*	2	Mid larva	ND
		s1132		Early larva	Mid larva
1	*let-71*	*s692*	2	Sterile (leaky)	Sterile (leaky)
1	*let-72*	*s52*	1	Late larva	ND
		s695	2	Mid larva	Mid larva
1	*let-73*	*s685*	2	Sterile	Sterile
1	*let-74*	*s697*	2	Late larva (leaky)	Late larva
1	*let-91*	*s678*	2	Mid larva	Mid larva
		s753	3	Late larva	Late larva
		s1165		ND	Mid larva
		s1185		Early larva	Egg to early larva
4	*let-92*	*s504*	2	Early larva	Early larva
		s677	2	Early larva	ND
7	*let-93*	*s734*	3	Mid larva	Mid larva
1	*let-96*	*s1112*		Mid larva	Mid larva
10	*let-97*	*s1121*		Early larva	Early larva
1	*let-98*	*s1117*		Late larva	Late larva
10	*let-99*	*s1201*		Maternal effect let	Maternal effect let
1	*let-100*	*s1160*		Early larva	Egg to early larva
1	*let-307*	*s1171*		Mid larva	Mid larva
1	*let-308*	*s1705*		Mid larva[e]	Mid larva
9	*let-309*	*s1115*		Late larva	Late larva
1	*let-311*	*s1195*		Late larva	Late larva
1	*let-312*	*s1234*		Late larva	Late larva
1	*lin-3*	*s751*	3	Late larva	Early larva
		s1263		Late larva	ND
		mDf7, sDf2, nDf27, sDf21		Egg	
		sDf7, sDf8, sDf9, sDf10		Early larva	

[a] Refer to Figure 2 for the location of each zone on the map.

[b] For information on the isolation, mapping and characterization of lethal mutations from reference 1, see ROGALSKI, MOERMAN and BAILLIE (1982) and MOERMAN (1980). For isolation and mapping of lethal mutations from reference 2, see ROGALSKI and BAILLIE (1985). The three reference 3 (DONATI 1985) lethal mutations were isolated in an identical manner to the screen described in this paper. The mutations listed without a reference are the 26 isolated in the screen described in this paper. The sources for the deficiencies listed here are indicated in MATERIALS AND METHODS.

[c] Lethal mutations described in references 1 and 2 are in *cis* with *unc-22(s7)*. Those described in reference 3 and in this paper are in *cis* with *unc-22(s7)* and *unc-31(e169)*. ND = not determined.

[d] The egg to early phenotype of some *let-x/sDf2* individuals indicates that a more precise developmental arrest stage was not determined (see MATERIALS AND METHODS).

[e] The phenotype of *let-308(s1705)* homozygotes and hemizygotes is with a background of *let-99(s1201)* in *cis* (see RESULTS).

noninteracting loci. In addition, the allele *s1165* only fails to complement *s1185* and neither of the other two alleles.

Lethal phenotypes of essential gene alleles and deficiencies: Table 2 is a list of all EMS-induced lethal mutations isolated to date in our laboratory in the *sDf2* region, as well as some deficiencies. The stages at which individuals homozygous for recessive lethal alleles in the *sDf2* region block development range from the early larva to the sterile adult. The homozygous (*let-x/let-x*) and, for most, the hemizygous (*let-x/sDf2*) phenotype for each lethal allele is indicated. The phenotypes for mutations isolated in this study were established as described in MATERIALS AND METHODS. Details about the previously isolated mutations can be found in the corresponding references listed in Table 2. Note that there are no egg lethal mutations in this region with the exception of some of the deficiencies. The phenotypes described as "leaky" are those where there are occasional fertile progeny. It should be noted here that the lethal phenotypes of some mutations may be due to a synthetic effect, since the mutations are all in *cis* with either *unc-22(s7)* or both *unc-22(s7)* and *unc-31(e169)*.

let-308(s1705) and *let-99(s1201)* were shown to be carried on the same chromosome by deficiency mapping. Homozygous individuals for this chromosome have a mid larval lethal phenotype. Individuals of

TABLE 3

Two-factor mapping data

Gene	Allele	Recombinants		Unc-31	Total	Position relative to *unc-22*[a]	Distance from *unc-22*[b]
		Unc-22Unc-31	Unc-22				
lin-3	*s751*	11		10	2769	left	0.60 (0.30–1.07)
let-93	*s734*		15	17	3661	right	0.62 (0.31–0.93)
let-96	*s1112*	19		8	3289	left	0.87 (0.48–1.26)
let-97	*s1121*		15	3	3925	right	0.57 (0.29–0.87)
let-98	*s1117*	17		8	3858	left	0.66 (0.35–0.97)
let-99	*s1201*		8	10	4714	right	0.25 (0.10–0.49)
let-100	*s1160*	17		21	4143	left	0.62 (0.33–0.91)
let-307	*s1171*	24		13	4615	left	0.78 (0.47–1.09)
let-308	*s1705*	18		12	3677	left	0.74 (0.40–1.08)
let-309	*s1115*		12	8	3693	right	0.49 (0.25–0.85)
let-311	*s1195*	11		12	3353	left	0.49 (0.25–0.88)
let-312	*s1234*	16		16	2442	left	0.99 (0.51–1.47)

[a] The hermaphrodites used for two-factor mapping of lethal mutations were of the genotype [*let-a(sx)*]*unc-22(s7)unc-31(e169)*/+ + +. The classes of recombinant progeny enable the placement of mutations to the right or left of *unc-22*. The presence of Unc-22 recombinants indicates that the mutation lies between *unc-22* and *unc-31* whereas Unc-22Unc-31 recombinants indicate a mutation that lies to the left of *unc-22*.

[b] Distance in map units or $100p$ where $p = 1 - [1 - (3U/\text{total})]^{1/2}$ and U = number of Unc-22 or Unc-22Unc-31 recombinants and total = number of individuals scored. The 95% confidence intervals are shown in parentheses and are calculated using the formula $1.96(Npq)^{1/2}$ where N = wild-type individuals (W), $p = U/W$ and $q = 1 - p$ or by using the table of STEVENS (1942).

the genotype *let-308(s1705) unc-22 let-99(s1201) unc-31/sDf21* are adult Unc-22 Unc-31s that lay dead eggs whereas *let-308(s1705) unc-22 let-99(s1201) unc-31/sDf2 or mDf7* individuals arrest development in the mid-larval stage. *let-99(s1201)*, when separated recombinationally from *let-308(s1705)*, also has a maternal egg lethal phenotype as a homozygote. The lethal phenotype of *let-308(s1705)* alone has not been determined.

DISCUSSION

We have recovered a set of 0.012 M EMS-induced recessive lethal mutations over the balancer *nT1(IV;V)*. The lethal mutation frequency is 8.7% in this 49-map-unit region. This frequency is comparable to that for the 44-map-unit *eT1(III;V)* region: 6.3% with the same EMS dose (R. ROSENBLUTH and R. JOHNSEN, personal communication). The *unc-22* region lethal mutations isolated in this screen were further characterized.

For this study, we define the *unc-22* region on linkage group IV as the interval spanned by the deficiency *sDf2*. This region is a minimum of 1.8 and a maximum of 2.5 map units. To date, 63 lethal and sterile mutations have been isolated which fall into 31 complementation groups. There are 23 essential genes and at least three nonessential genes (*dpy-20*, *him-8*, and *mec-3*) to the left of *unc-22* and 8 essential genes and one non-essential gene (*dpy-26*) to the right (see EDGLEY and RIDDLE 1987 for information on nonessential genes). The right breakpoint of *sDf2* is not clearly defined (0.6–1.3 map units from *unc-22*) and the interval to the right of *unc-22* is gene sparse relative to the left interval.

The apparent clustering of genes to the left of *unc-22*, as noted previously (ROGALSKI, MOERMAN and BAILLIE 1982; ROGALSKI and BAILLIE 1985), could result from the following: (1) a lower frequency of recombination on the left of *unc-22* relative to the right, (2) a higher number of mutationally silent genes to the right of *unc-22*, and (3) a physical clustering of genes, with more genes per length of DNA in the left interval than in the right. The clustering of genes on the recombination map of *C. elegans* was first noted by BRENNER (1974) for the nonessential genes on each linkage group. Essential gene clustering was subsequently found to coincide with the clustering of nonessential genes on LGI (ROSE and BAILLIE 1980; HOWELL *et al.* 1987), LGII (SIGURDSON, SPANIER and HERMAN 1984) and LGV (ROSENBLUTH *et al.* 1988). Recent molecular work on the LGIII gene cluster containing *lin-12* shows that the clustering in that region appears to be due to recombinational suppression and not to an uneven distribution of genes on the physical map (GREENWALD *et al.* 1987).

The *unc-22* region has now been divided into ten zones which are defined by the breakpoints of nine overlapping deficiencies including *sDf2*. Deficiency breakpoints serve to unambiguously order complementation groups on the genetic map. It is apparent from Figure 2 that more deficiency breakpoints are required, particularly in zone 1, to order the complementation groups.

Using the truncated Poisson formula of MENEELY

and HERMAN (1979) and the number of alleles for each essential gene identified, we predict that there is a minimum of 48 essential genes in the *sDf2* region. The Poisson distribution inherently carries the assumption that all of the essential genes are equally mutable. Work with *Drosophila melanogaster* has shown that this assumption cannot be made when attempting to accurately estimate gene number (HILLIKER, CHOVNICK and CLARK 1981; LEFEVRE and WATKINS 1986). Some of the genes are clearly "hot spots" for EMS mutagenesis (see Table 2). This is particularly true for *let-56* which has eight alleles when the mean allele frequency is about one. It must therefore be emphasized that our estimate is only a minimum one. Another factor that can distort a gene number estimate, making it too high, is interallelic complementation. With the addition of more deficiency breakpoints and more alleles of each gene, the presence of complementing alleles will be revealed if they do exist. *let-91* is the one locus in the *sDf2* region that shows a complex complementation pattern. For the purpose of calculating essential gene number, it has been considered a single locus with four alleles.

By comparing the number of lethal mutations obtained in the *nT1(IV;V)* region to that in the *sDf2* region, one can arrive at an estimate of the number of essential genes in the *nT1* balanced region which can then be extended to an estimate for the genome. This calculation is not dependent on a uniform distribution of essential genes in the *nT1* balanced region, but it again depends on the assumption that the genes are equally mutable. The estimated number of essential genes in the *sDf2* region is 48, and 25 chromosomes carrying lethal mutations in the *sDf2* region were isolated in the *nT1* screen. In parallel, since 294 lethal mutations were isolated in the *nT1* balanced region, there is a minimum of 565 essential genes in the 49 map unit *nT1* region, which is approximately one sixth of the genome. This leads to a minimum estimate of approximately 3500 essential genes for the *C. elegans* genome, which is not far from the original estimate of 2000 essential genes (BRENNER 1974). There is a problem in extending the essential gene estimate for the *nT1* region to one for the entire genome. The *nT1* region includes two of the five autosomal gene clusters while it only includes two of the ten autosomal arms. Therefore, the average essential gene density for the *nT1* region is probably higher than that for the overall genome. For this reason, our essential gene number estimate for the genome may be an overestimate.

The lethal phenotypes of EMS induced mutations in the *sDf2* region range from early larvae to sterile adults. Previously (ROGALSKI, MOERMAN and BAILLIE 1982), it was noted that no lethal allele in the *sDf2* region had an egg lethal phenotype. With the addition of 26 mutations from this study to the previously isolated 37, there are still no egg lethal alleles. However, *sDf2* homozygotes arrest development in the egg, while *sDf7*, *sDf8*, *sDf9* and *sDf10* homozygotes arrest as early larvae (Figure 2 and Table 2). This indicates that there may be an as yet unidentified essential gene with an egg lethal mutant phenotype in either zone 1, 8, 9 or 10. Since we have calculated that not all essential genes have been identified, this essential gene could be identified in the future. Alternatively, it is possible that one of the 31 identified essential genes has an egg lethal null phenotype and that the null allele has not yet been isolated. A third explanation for the egg lethal phenotype of *sDf2* homozygotes is that it is due to a synthetic effect of deleting a number of essential genes. The possibility that there is a second mutation on the *sDf2* chromosome causing the egg lethal phenotype is small, since *sDf2* was originally maintained in an unbalanced strain.

Of the 63 *sDf2* region lethal mutations, 49 were analyzed with respect to their phenotypes in both *let-x/let-x* and *let-x/sDf2* individuals. In 11 cases, the observable stage of developmental arrest varies from homozygote to hemizygote. Nine lethal mutations exhibit a more severe phenotype in the hemizygote, indicating that these are hypomorphic mutations. Two mutations, *let-66(s176)* and *let-70(s1132)*, exhibit a mid larval stage of arrest in the hemizygote and an early larval one in the homozygote. Either *s176* and *s1132* are hypermorphic (or antimorphic) or there are second mutations in the *nT1* balanced region in the strains carrying these mutations.

It may be that more of the lethal alleles in Table 2 have varying phenotypes in homozygotes versus hemizygotes, since the method we use to classify lethal phenotypes, namely by length of the nematode, is approximate. Any fine distinctions between phenotypes would probably be missed. What we present here is a guideline of lethal phenotypes for each essential gene. This guideline gives us an indication of the stages of development in which an essential gene product is required. We make the assumption that the mutant phenotype of an essential gene is an indication of its time of expression in development and that the gene product is required before or at the time of developmental arrest.

The continued analysis of the *unc-22* region in our laboratory will allow a correlation of the genetic and physical maps. This can be accomplished in several ways. First, genetically mapped deficiency breakpoints can serve to identify the extent of essential genes on the physical map, as was accomplished in the *rosy-ace* region of *D. melanogaster* (GAUSZ *et al.* 1986). Second, isolation of lethal alleles in a mutator strain that has high mobility of the transposable element Tc1 (MOERMAN and WATERSTON 1984) allows tagging and cloning of essential genes (K. MCKIM,

D. CLARK, R. JOHNSEN and D. BAILLIE, unpublished results). Third, cosmids associated with the genetically mapped RFLDs can be used to rescue mutant strains through germ line transformation (FIRE 1986).

Correlation of the genetic and physical maps of the *unc-22* region will allow us to determine the density of essential genes per kilobase. We can also determine whether there is variability in essential gene density on either side of *unc-22* on the physical map as was found on the genetic map. Additionally, the identification of coding elements, by finding regions of sequence identity between *C. elegans* and *Caenorhabditis briggsae* (S. PRASAD and D. BAILLIE, unpublished results), will enable us to determine the number of mutationally silent genes in the *unc-22* region. Eventually, the association of coding elements with the genetic map and the study of their expression at the molecular level will reveal much about genome organization in *C. elegans*.

The authors wish to thank RAJA ROSENBLUTH, ANN ROSE, BOB JOHNSEN, KIM MCKIM and LINDA ADDISON for helpful discussion throughout this work and M. R. ROGALSKI for technical assistance. We also give thanks to the reviewers for their useful and helpful comments on the manuscript. This research was supported by Natural Sciences and Engineering Research Council of Canada and Muscular Dystrophy Association of Canada grants to D.L.B. and a Muscular Dystrophy Association of Canada predoctoral fellowship to D.V.C.

LITERATURE CITED

BAILLIE, D. L., K. A. BECKENBACH and A. M. ROSE, 1985 Cloning within the *unc-43* to *unc-31* interval (linkage group IV) of the *Caenorhabditis elegans* genome using Tc1 linkage selection. Can. J. Genet. Cytol. **27:** 457–466.

BRENNER, S., 1974 The genetics of *Caenorhabditis elegans*. Genetics **77:** 71–94.

COULSON, A., J. SULSTON, S. BRENNER and J. KARN, 1986 Toward a physical map of the genome of the nematode *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA **83:** 7821–7825.

DONATI, L. A. M., 1985 A genetic analysis of the right arm of linkage group IV of *Caenorhabditis elegans*, with emphasis on the *sDf2* region. M.Sc. Thesis, Simon Fraser University, Burnaby, B.C., Canada.

EDGLEY, M. L., and D. L. RIDDLE, 1987 *Caenorhabditis elegans*. In: *Genetic Maps 1987*, Vol. 4, Edited by S. J. O'BRIEN. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

FERGUSON, E. L., and H. R. HORVITZ, 1985 Identification and characterization of 22 genes that affect the vulval cell lineages of the nematode *Caenorhabditis elegans*. Genetics **110:** 17–72.

FIRE, A., 1986 Integrative transformation of *Caenorhabditis elegans*. EMBO J **5:** 2673–2680.

GAUSZ, J., L. M. C. HALL, A. SPIERER and P. SPIERER, 1986 Molecular genetics of the *rosy-ace* region of *Drosophila melanogaster*. Genetics **112:** 65–78.

GREENWALD, I., A. COULSON, J. SULSTON and J. PRIESS, 1987 Correlation of the physical and genetic maps in the *lin-12* region of *Caenorhabditis elegans*. Nucleic Acids Res. **15:** 2295–2307.

HILLIKER, A. J., A. CHOVNICK and S. H. CLARK, 1981 The relative mutabilities of vital genes in *D. melanogaster*. Drosophila Inform. Serv. **56:** 64–65.

HORVITZ, H. R., S. BRENNER, J. HODGKIN and R. K. HERMAN, 1979 A uniform genetic nomenclature for the nematode *Caenorhabditis elegans*. Mol. Gen. Genet. **175:** 129–133.

HOWELL, A. M., S. G. GILMOUR, R. A. MANCEBO and A. M. ROSE, 1987 Genetic analysis of a large autosomal region in *Caenorhabditis elegans* by the use of a free duplication. Genet. Res. **49:** 207–213.

LEFEVRE, G., and W. WATKINS, 1986 The question of total gene number in *Drosophila melanogaster*. Genetics **113:** 869–895.

MENEELY, P. M., and R. K. HERMAN, 1979 Lethals, steriles, and deficiencies in a region of the X-chromosome of *Caenorhabditis elegans*. Genetics **92:** 99–115.

MENEELY, P. M., and R. K. HERMAN, 1981 Suppression and function of X-linked lethal and sterile mutations in *C. elegans*. Genetics **97:** 65–84.

MOERMAN, D. G., 1980 A genetic analysis of the *unc-22* region in *Caenorhabditis elegans*. Ph.D. thesis, Simon Fraser University, Burnaby, B.C., Canada.

MOERMAN, D. G., and D. L. BAILLIE, 1979 Genetic organization in *Caenorhabditis elegans*: Fine structure analysis of the *unc-22* gene. Genetics **91:** 95–103.

MOERMAN, D. G., and D. L. BAILLIE, 1981 Formaldehyde mutagenesis in *Caenorhabditis elegans*. Mut. Res. **80:** 273–279.

MOERMAN, D. G., and R. H. WATERSTON, 1984 Spontaneous unstable *unc-22 IV* mutations in *C. elegans* var. Bergerac. Genetics **108:** 859–877.

MOERMAN, D. G., G. M. BENIAN and R. H. WATERSTON, 1986 Molecular cloning of the muscle gene *unc-22* in *C. elegans* by *Tc1* transposon tagging. Proc. Natl. Acad. Sci. USA **83:** 2579–2583.

ROGALSKI, T. M., and D. L. BAILLIE, 1985 Genetic organization of the *unc-22 IV* gene and the adjacent region in *Caenorhabditis elegans*. Mol. Gen. Genet. **201:** 409–414.

ROGALSKI, T. M., and D. L. RIDDLE, 1988 A *Caenorhabditis elegans* RNA polymerase gene, *ama-1 IV*, and nearby essential genes. Genetics **118:** 61–74.

ROGALSKI, T. M., D. G. MOERMAN and D. L. BAILLIE, 1982 Essential genes and deficiencies in the *unc-22 IV* region of *Caenorhabditis elegans*. Genetics **102:** 725–736.

ROSE, A. M., and D. L. BAILLIE, 1979 The effect of temperature and parental age on recombination and nondisjunction in *Caenorhabditis elegans*. Genetics **92:** 409–418.

ROSE, A. M., and D. L. BAILLIE, 1980 Genetic organization of the region around *unc-15(I)*, a gene affecting paramyosin in *Caenorhabditis elegans*. Genetics **96:** 639–648.

ROSENBLUTH, R. E., and D. L. BAILLIE, 1981 The genetic analysis of a reciprocal translocation, *eT1(III;V)*, in *Caenorhabditis elegans*. Genetics **99:** 415–428.

ROSENBLUTH, R. E., T. M. ROGALSKI, R. C. JOHNSEN, L. M. ADDISON and D. L. BAILLIE, 1988 Genomic organization in *Caenorhabditis elegans*: deficiency mapping on linkage group V (left) Genet. Res. In press.

SIGURDSON, D. C., G. J. SPANIER and R. K. HERMAN, 1984 *Caenorhabditis elegans* deficiency mapping. Genetics **108:** 331–345.

STEVENS, W. L., 1942 Accuracy of mutation rates. J. Genet. **43:** 301–307.

Communicating editor: V. G. FINNERTY