GENETIC ANALYSIS OF THE LEFT HALF OF LINKAGE GROUP V IN *CAENORHABDITIS ELEGANS*

by

Robert Craig Johnsen

B.Sc., Simon Fraser University, 1986

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in the Department

of

Biological Sciences

Simon Fraser University

October 1990

APPROVAL

Name: ROBERT CRAIG JOHNSEN

Degree: Doctor of Philosophy

Title of Thesis:

GENETIC ANALYSIS OF THE LEFT HALF OF LINKAGE GROUP V IN CAENORHABDITIS ELEGANS

Examining Committee:

Chairman: Dr. M.J. Smith, Professor

Dr. D.L. Baillie, Professor, Senior Supervisor, Dept. of Biological Sciences, SFU

Dr. A.T. Beckenbach, Associate Professor, Dept. of Biological Sciences, SFU

Dr. A.M. Rose, Associate Professor, Genetics, UBC

Dr. D.G. Holm, Professor, Genetics, UBC

R.E. Rosenbluth, Research Technician, Dept. of Biological Sciences, SFU

Dr. B.P. Brandhorst, Professor, Dept. of Biological Sciences, SFU, Public Examiner

Dr. D.G. Moerman, Assistant Professor, Genetics, UBC, Public Examiner

Dr. I.S. Greenwald, Assistant Professor, Dept. of Biology, Princeton University, Princeton, New Jersey, External Examiner

Date Approved October 31 1990

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

GENETIC ANALYSIS OF THE LEFT HALF OF LINKAGE GROUP V IN CAENORHABDITIS ELEGANS

Author:

(signature)

ROBERT JOHNSEN

(name)

Oct 17/90

(date)

ABSTRACT

This is a study of genome organization. Two questions that I address are: Is there logic to the recombinational clustering of genes? Have selective pressures caused genes of similar function or time of action to be juxtaposed? Before these questions can be answered, large genomic regions must be analyzed in detail. This is the largest genomic region that has been studied for genetic organization in a higher eukaryote. The region is LGV(left) which is balanced by *eT1(III;V)*. LGV(left) comprises approximately 7% of the recombinational distance in *Caenorhabditis elegans*. Mutagens that induce rearrangements, mobilization of transposons (Tc1) and point mutations were used to generate the mutations needed for the analysis. LGV(left) was divided into zones using overlapping deficiencies. Recessive lethal point mutations were mapped to genes by appropriate complementation tests. The set of overlapping deficiencies were used to analyze recombination. Mobilization of Tc1 induced mutations that were used to assess the types of mutations induced by transposons also aided in cloning genes. The mutations will also be useful as anchor points for the alignment of the physical and genetic maps.

10,900 EMS treated F_1's were screened, yielding 194 lethal mutations on LGV(left). The analysis of 166 of the lethals resulted in the identification of one deficiency and alleles of 74 essential genes including 38 new genes. A total of 101 genes on LGV(left) with recessive lethal alleles have now been identified. The average hit frequency per gene is 1.25; there are a minimum of 147 essential genes on LGV(left) and 3500 in the genome. LGV(left) has two gene clusters separated by a high recombination/gene sparse region. Two genes in the sparse region have interesting characteristics: *let-330* is the largest EMS target on LGV(left) with twice as many alleles as the next largest gene; *lin-40* (a complex locus that affects the switch from spermatogenesis to oogenesis) and the region near it are major targets for Tc1 mobilization induced mutagenesis. In addition, a

region near the left tip of LGV was identified as important for recombination and a model of chromosome pairing for recombination has been erected.

This work has led to insights into genomic organization and has also defined a large genetic region that can be correlated with the physical map and ultimately with the DNA sequence.

This work is dedicated to my parents.

ACKNOWLEDGEMENTS

I would to thank my senior supervisor Dr. David L. Baillie for his enthusiasm, insightful suggestions and encouragement. I would also like to thank the members of my supervisory committee, Drs A.T. Beckenbach, D.G. Holm, A.M. Rose and especially R.E. Rosenbluth for many helpful discussions in the preparation of this thesis. Many thanks to H. Stewart for maintaining stocks and technical assistance; M. Rogalski for technical assistance; K. Beckenbach for help with the molecular work; and Dr. P. Sternberg for help in identifying gonadal abnormalities. I would also like to thank D. Clark, M. Wakarchuk, M. Marra, B. Barbazuk and my friends and family for their support.

This work was financially supported by a Medical Research Council of Canada Studentship.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

GENERAL INTRODUCTION

This thesis is a genetic study of the organization of a large region of the genome in the nematode *Caenorhabditis elegans*.

One of the most interesting problems in biology has to do with the relationship of genome structure and genome function. To uncover the underlying principles of genomic organization, genetic and molecular analysis must be done on large chromosomal regions in highly diverged species. I have chosen to work with *C. elegans* because it provides an excellent counterpoint to the work done in *Drosophila melanogaster*, *Saccharomyces cerevisiae* and *Mus musculus*.

D. melanogaster has one of the best characterized eukaryotic genomes. The 165 million base pairs of DNA in *D. melanogaster's* haploid genome consists of three classes of sequences. Crain *et al.* (1976) used reassociation kinetics to show that 60% of the genome is single copy sequence, 25% is highly repeated sequence and the remainder is heterogeneous-moderately repeated DNA. The highly repeated DNA, including the foldback elements (Schmid *et al.* 1975), are located in the heterochromatic regions of all the chromosomes (Peacock *et al.* 1974). The moderately repeated DNA consists of transposons, and reiterated gene families such as the histones, 5S and rRNA genes.

Many related genes are known to exist as clusters on the chromosomes. Wright (1987) divided gene clusters into four classes: 1) reiterated genes, 2) gene complexes, 3) functionally and evolutionarily (structurally) related genes and 4) functionally but not evolutionarily (structurally) related genes.

The homeotic bithorax and antennapedia complexes are examples of gene complexes. They are both responsible for specifying some of the spatial organization of Drosophila. Some of the clusters of functionally and structurally related genes may be the result of recent gene duplications but others may exist for regulatory reasons. Two

examples of functionally and structurally related clustered genes are the glue genes (Martin and Meyerowitz 1988) and the yolk protein genes (Barnett *et al.* 1980). Two examples of functionally but not structurally related genes are: 1) the dopa decarboxylase structure genes (*Ddc*) where 13 genes which affect the formation, sclerotisation or pigmentation of the cuticle (Wright *et al.* 1984) occur in an 8-12 polytene band region; 2) the myofibrillon contractile protein genes and actin 88F whose products only accumulate in the indirect flight muscles of the thorax (Karlik *et al.* 1984).

On genomic recombination maps genes are not distributed randomly along chromosomes but show clustering [for a Drosophila review see Baker *et al.* 1976; for yeast see Lambie and Roeder 1986; for *C. elegans* see the Genomic Maps of Edgley and Riddle (1990)]. Results from identifying the number of coding elements in regions of known length both physically and recombinationally suggest that gene spacing along the chromosome and recombinational distances do not have a strong correlation (Greenwald *et al.* 1987; Prasad and Baillie 1989; Starr *et al.* 1989). These results imply that not only are there genes that are physically clustered for functional reasons but there are genes physically close but widely separated recombinationally and others that are recombinationally close but physically well separated. The reasons for these phenomena are not yet clear.

C. elegans has many advantages that make it an excellent model system for genetic analysis (Brenner 1974). It has all the basic characteristics of higher eukaryotes: it has a nervous system, muscle, gut, integument and a complex reproductive system. *C. elegans* is small and requires little space to maintain and store. Maintenance is relatively easy because strains can be stored in liquid nitrogen. Freezing in liquid nitrogen not only precludes the necessity for stock maintenance but, even more importantly, it probably slows the accumulation of spontaneous mutations in balanced regions. Mutations may arise through replicational error, chemical mutagenesis or background radiation; freezing

the nematodes in liquid nitrogen effectively eliminates the first two mechanisms and thus protects the balanced region from further genetic insult. In addition, larvae of unfrozen strains can enter a dormant "dauer larva" stage under conditions of overcrowding and limited food. Dauers can survive for several months at room temperature. *C. elegans* has few chromosomes (5AA,XX); (a simplified genetic map is presented in Figure 1), a short generation time (3.5 days at 20C) and produces a large number of easily visible eggs (approximately 330 per wild-type hermaphrodite). *C. elegans*' common form of reproduction is hermaphroditic self-fertilization; this eliminates the necessity of brother-sister matings in order to obtain homozygosity of new mutations in the F_2 generation, but males exist and thus allow standard genetic analysis.

C. elegans strains are routinely maintained, for indefinitely long periods, in liquid nitrogen and the entire history of most strains should be traceable to the original N2 (defining *C. elegans*' var. Bristol wild-type) hermaphrodite isolated by Brenner (1974) for genetic studies. Therefore, the isolation of a large number of mutations has the long term benefit of providing well defined strains for future genetic and molecular analysis.

Various aspects of *C. elegans* have been intensively studied by many investigators with the ultimate goal of completely defining the organism. In order to completely define an organism it is necessary to identify all of its genes. So far, over 800 genes (both essential and non-essential) have been identified (Edgley and Riddle 1990). Minimum essential gene number estimates ranged from 2000 (Brenner 1974; Meneely and Herman 1981) to 3500 (Clark *et al.* 1988) or 4000 (Howell 1989). Attempts to map all the essential genes in some regions of *C. elegans* have shown that essential genes probably comprise 80-90% of all genes (for the region around *unc-15(I)*, see Rose and Baillie 1980; Howell 1989; for Linkage Group (LG)II see Sigurdson *et al.* 1984; for the region around *unc-22(IV)* see Rogalski *et al.* 1982; Rogalski and Baillie 1985; Clark 1990). The complete cell lineage of *C. elegans* has been elucidated (Sulston and Horvitz 1977; Sulston *et al.*

1983), a genomic cosmid-Yac contig map is nearing completion (A. Coulson, J.E. Sulston pers. comm.) and the sequencing of the *C. elegans* genome will soon begin (A. Coulson, J.E. Sulston and R.H. Waterston pers. comm.).

The identification of the majority of essential genes in a large genomic region of *C. elegans* would have many benefits including: 1) providing a more accurate calculation of the minimum number of essential genes necessary to build a higher organism. 2) aiding in the correlation of a large region of the sequenced physical map and the genetic map. The sequence data will allow for the identification of all genes, but may not yield much information about their biological functions. The existence of strains containing mutations in most of those genes should supply functional information. 3) providing greater insight into genomic organization. Genes are nonrandomly distributed along recombination maps of the chromosomes (Brenner 1974; see Genomic Maps by Edgley and Riddle 1990). The study of the distribution of blocking stages and/or null phenotypes of genes in a large region may lead to a rationale for the recombinationally nonrandom gene distribution. 4) establishing a set of landmarks useful for locating structural features along the chromosome such as pairing points. 5) supplying new alleles of some of the intensively studied genes and identifying and partially characterizing previously unknown genes. 6) defining a large genomic region which provides an excellent system for testing mutagens in an eukaryote (Rosenbluth *et al.* 1983).

In this thesis I have attempted to identify the majority of essential genes in the 23 map units (m.u.) of the left half of LGV which is recombinationally balanced by the reciprocal translocation *eT1(III;V)* (*eT1*) (Rosenbluth and Baillie 1981). LGV(left) contains 7% of the recombinational distance in *C. elegans'* genome. LGV(left) has been subdivided into 22 zones by sets of overlapping deficiencies and 1 duplication (Rosenbluth *et al.* 1988). Previous workers have identified 17 putative non-essential genes (under laboratory

conditions) and 59 essential genes on LGV(left) (Rosenbluth *et al.* 1988; Edgley and Riddle 1990; Stewart *et al.* 1990).

Since *eT1* balances not only LGV(left) but also 20 m.u. on LGIII(right) (*i.e.* a total of 43 m.u.), it provides excellent system for testing mutagens. By screening for recessive lethal mutations in the balanced regions, the effect of various mutagens on LGIII(right) and LGV(left) have been compared quantitively. The effect of mutagens on LGV(left) has been analyzed in more detail. The mutagens used include: gamma radiation and ethyl methanesulfonate (EMS) (Rosenbluth *et al.* 1983; Rosenbluth *et al.* 1988); ultraviolet radiation (Stewart *et al.* 1990); high and low LET (linear energy transfer) ionizing radiation (Nelson *et al.* 1989).

In Section 1 of this thesis, the number of zones that subdivide LGV(left) was increased to facilitate mapping essential genes on LGV(left). Since Moerman and Baillie (1981) showed that formaldehyde induces small deletions in *C. elegans*, formaldehyde was chosen to induce a new set of deficiencies in LGV(left). A formaldehyde dose-response curve was developed and a number of formaldehyde induced mutations including some deficiencies were isolated and characterized.

In Section 2, mutations induced on LGV(left) by the mobilization of the transposon Tc1 were analyzed. Tc1 is activated by *mut-4(I)* (Mori *et al.* 1988b) and is an excellent molecular tag for specific genes (see Greenwald 1985; Moerman *et al.* 1986 for examples; for a review on Tc1 see Moerman and Waterston 1989). In attempting to Tc1 tag a set of genes at random on LGV(left) (to be used as anchor points for correlating the genetic and molecular maps), it became apparent that Tc1 showed some strong mutational biases: a gene, *lin-40*, and the region around *lin-40* are very susceptible to Tc1 induced mutagenesis.

In Section 3, a combined study by R.E. Rosenbluth and myself on the effect on recombination in deficiency heterozygotes is presented. In the course of analyzing the

deficiencies that sub-divide LGV(left) into zones it was noticed that some deficiencies suppress recombination. Upon analyzing these effects further we established that the deficiencies fell into two classes of suppressors (major and minor); all the internal deficiencies studied were major inhibitors. The major inhibitors show polarity of inhibition, with effectively no inhibition on their left but nearly complete inhibition on their right. In some cases recombination is inhibited by 95% over a 16 m.u. distance. From the effects of the deficiency heterozygotes on recombination we developed a model for pairing for recombination in *C. elegans*.

In Section 4, the screening and analysis of EMS induced lethal mutations on LGV(left) is presented. In preliminary studies, I, in collaboration with R.E. Rosenbluth, T.M. Rogalski, L.M. Addison and D.L. Baillie (Rosenbluth *et al.* 1988), and Stewart *et al.* (1990), established an initial set of lethals. To increase the number of identified essential genes on LGV(left) to approximately 70% of the total number of essential genes, 10,900 *eT1* balanced chromosomes treated with 0.012M EMS were screened. From the screen, 751 recessive lethal [including sterile adult and some maternal effect lethal (*mel*)] mutations in or close to the balanced regions were isolated. The number of lethals that mapped to either of the two *eT1* balanced chromosomes were compared. Complementation tests were used to map the lethals on LGV(left) against a set of deficiencies that divide LGV(left) into zones, and once the lethals were mapped to zones complementation tests were used to map them against appropriate genes to discern if the lethals identified new genes or were alleles of known genes. The approximate blocking stage for each lethal at 20C was noted and the mutations were tested to see if any were rescuable at 15C or 25C. The sterile adults and Mels were tested for male rescuability.

In Section 5, two genes, *lin-40* and *let-336*, both of which cause large brood sizes as mutant heterozygotes, were analyzed. The analysis of deficiencies on LGV(left) led to a second interesting result (the first being recombination suppression due to heterozygous

deficiencies): while scoring the progeny from two-factor mapping experiments, some anomalously large self-progeny brood sizes were noted for hermaphrodites containing some heterozygous deficiencies, *i.e.* hermaphrodites containing deficiencies that delete *lin-40* or *let-336* had unusually large brood sizes. This indicated that *lin-40* and *let-336* are probably involved in the switch from spermatogenesis to oogenesis.

lin-40 has other interesting characteristics; it is a complex locus with at least five classes of complementing alleles, and it is highly susceptible to Tc1 induced mutagenesis. The cloning of *lin-40* using an allele putatively tagged with Tc1 was attempted. The resultant clone was sent to Cambridge, England for positioning on the physical map of *C. elegans* constructed by Coulson *et al.* (1986; 1988). The position of the clone on the physical map (close to *ges-1*) was surprisingly far to the left of the position expected from the location of *lin-40* on the genetic map.

Figure 1

Partial genetic map of *C. elegans*.

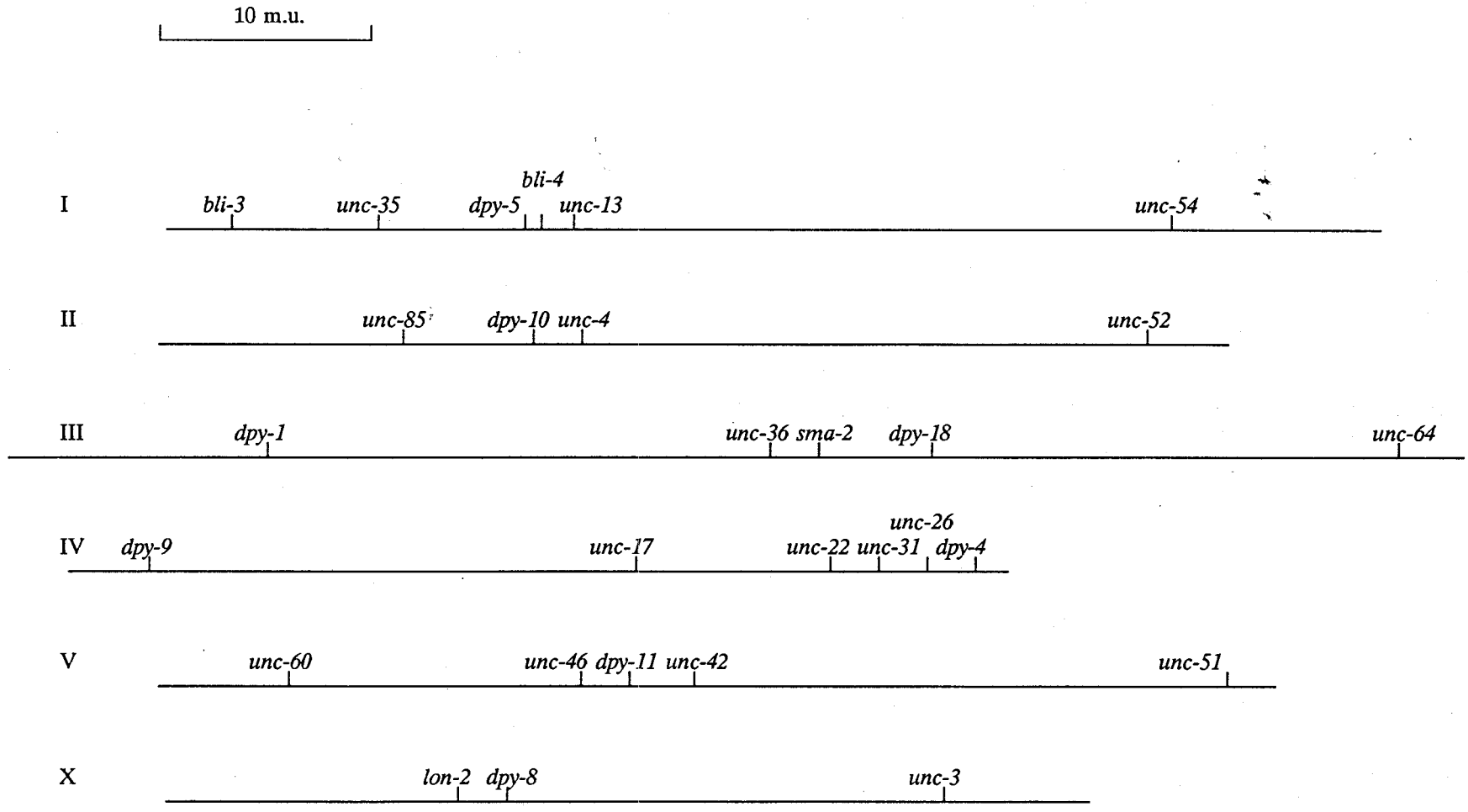
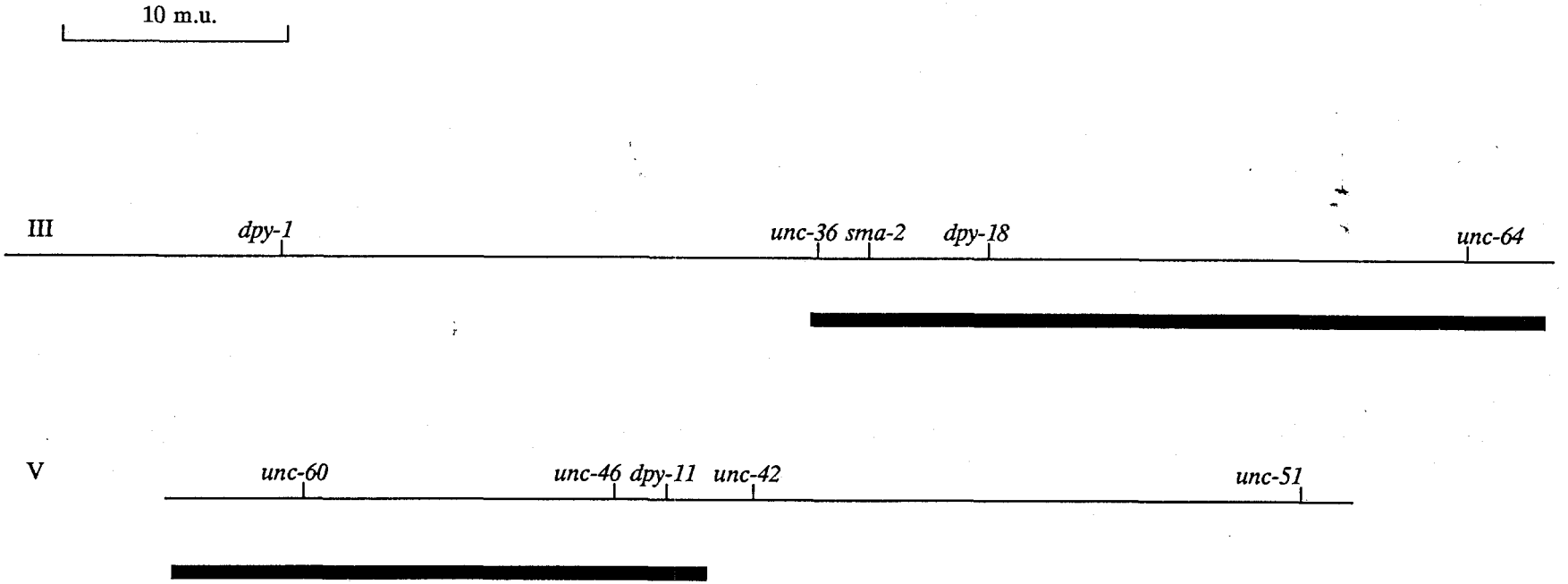

10 m.u.
I
bli-3
unc-35
bli-4
dpy-5
unc-13
unc-54
II
unc-85
dpy-10
unc-4
unc-52
III
dpy-1
unc-36
sma-2
dpy-18
unc-64
IV
dpy-9
unc-17
unc-26
unc-22
unc-31
dpy-4
V
unc-60
unc-46
dpy-11
unc-42
unc-51
X
lon-2
dpy-8
unc-3

Figure 2

LGIII and LGV genetic map with the regions balanced by the reciprocal translocation *eT1* underlined.


10 m.u.
III
dpy-1
unc-36
sma-2
dpy-18
unc-64
V
unc-60
unc-46
dpy-11
unc-42
unc-51

6 B

GENERAL MATERIALS AND METHODS

(i) *General*

The nomenclature follows the uniform system adopted for *C. elegans* (Horvitz *et al.* 1979). The nematodes were cultured in Petri dishes on a simple agar nematode growth medium streaked with *Escherichia coli* (OP50). For details of this as well as procedures for observing and handling the worms see Brenner (1974).

(ii) *Mutations*

The following *C. elegans* (var. Bristol) strains were obtained from the MRC, Cambridge, England, from the Caenorhabditis Genetics Center at the University of Missouri, Columbia, MO. or as cited: N2(wild-type), *unc-35(e259)I, dpy-5(e61)I, unc-52(e444)II, dpy-18(e364)III, unc-64(e246)III, dpy-4(e1166)IV, unc-26(e345)IV, dpy-11(e224)V, dpy-11(s287)eT1(III), unc-60(s1331)eT1(III), emb-29(g52)V, lin-40(e2173)V* [isolated by S.W. Emmons, was supplied by J. Hodgkin (MRC Cambridge)], *rol-3(e754)V, unc-23(e25)V, unc-34(e315* and *e566)V unc-42(e270)V, unc-46(e177)V, unc-60(e677 and m35)V, unc-62(s472)V, unc-68(e540)V*; *unc-76(911)V*; *ama-2(m323)* [Rogalski *et al.* (1988) from D.L. Riddle's laboratory (Columbia, MO.)]; *eT1(III;V)* [in which a reciprocal exchange had occurred between LGIII and LGV near the center of each chromosome (Rosenbluth and Baillie 1981). The two translocation chromosomes are *eT1(III)* consisting of LGIII(left)LGV(left), and *eT1(V)* consisting of LGIII(right)LGV(right)]. The deficiencies *mDf1* and *mDf3* (Brown 1984) were from D. L. Riddle's laboratory (Columbia, MO). *nDf32* (Park and Horvitz 1986b), *nDf18* and *nDf31* originated in R. H. Horvitz's laboratory (MIT). The translocation *nT1(IV;V)* (Ferguson and Horvitz 1985) has been shown to balance LGV from *unc-60* to *unc-76* (Ferguson and Horvitz 1985; Clark *et al.* 1988). All mutations denoted with the *s* prefix arose in this laboratory.

(iii) *Characteristics of eT1(III;V)*

eT1(III;V) is a reciprocal translocation that recombinationally balances the right half of LGIII and the left half of LGV. The balanced region of each chromosome is approximately the same size recombinationlly. A total of about 43 m.u. (15% of the genome) is balanced (Rosenbluth and Baillie 1981). Ten sixteenths of the progeny of *+/eT1(III);+/eT1(V)* hermaphrodites are aneuploids and block early in development, therefore LGIII(right) and LGV(left) are effectively pseudolinked. Homozygous *eT1* nematodes have an *unc-36* phenotype (*eT1* probably breaks in *unc-36 III).*

(iv) *BC2200: the test strain for formaldehyde and EMS mutagenesis*

Hermaphrodites from a homozygous *dpy-18;unc-46* strain were mated to *dpy-18/eT1;unc-46/eT1* males. Wild-type male offspring were mated to homozygous *eT1* hermaphrodites. One individual wild-type hermaphrodite from the progeny of the latter cross was used to establish the strain BC2200.

The genotype is *dpy-18(e364)/eT1(III);unc-46(e177)/eT1(V).* Due to the early death of the aneuploids, hermaphrodites with this genotype give rise to 1 Dpy Unc : 4 wild-type : 1 Unc-36 progeny. Since the markers *dpy-18(III)* and *unc-46(V)* are in the balanced regions, they are pseudolinked.

(v) *Definition of lethal, sterile and mel (maternal effect lethal)*

I use the word lethal in a population genetics sense. A recessive lethal mutation is one that when homozygous causes the worm not to mature (egg lethal, early, mid or late larval lethal); or to mature but not produce fertilized eggs (sterile); or to produce fertilized eggs that do not mature due to *mel* mutations. Second generation mels are also considered to be lethals.

(vi) *Screening protocol for formaldehyde and EMS induced lethal mutations*

Single wild-type F_1 heterozygotes from mutagenized P_0s (see separate sections for mutagenesis), were placed on individual Petri plates (10x60mm). From any one P_0

hermaphrodite plate, all F_1 heterozygotes were picked. The progeny of these F_1s were screened for the presence of normal appearing gravid Dpy Unc worms. All such F_1 lines were discarded. The absence of mature Dpy Uncs indicated the presence of a lethal within the screening region. From every plate having no fertile adult Dpy Uncs, a putative lethal bearing wild-type heterozygote was picked and used to establish a stock.

(vii) *Mapping formaldehyde and EMS induced mutations to LGIII or LGV*

Replacing the balancer (*eT1*) chromosomes with wild-types allows one to map the lethal to a chromosome and to calculate its distance from the appropriate marker. The lethal bearing strains were crossed to N2 males and several wild-type L4 F_1 hermaphrodites from each strain were picked. Any F_1s that gave Unc-36 progeny were discarded. In the formaldehyde experiments, for one or two of the F_1s, all the F_2s were scored. The lethal was assigned to LGIII(right) if the number of fertile F_2 Dpy-18s was significantly less than one third the number of wild-types; if the Unc-46 : wild-type ratio was significantly less than one third, the lethal was assigned to LGV(left). Strains with mutations on both chromosomes were not analyzed further. In the EMS experiments, if the progeny of lethal bearing *dpy-18/+;unc-46/+* were in an approximate 3 wild-type : 1 Dpy-18 ratio then the lethal mutation was on LGV(left) and the strain was retained, otherwise the strain was discarded (the strains with mutations on both LGIII and LGV were discarded).

(viii) *Mapping to location on LGV(left)*

LGV(left) has been divided into zones by a set of rearrangement breakpoints (Rosenbluth *et al.* 1988). A subset of these deficiencies (*sDf26*, *sDf30*, *sDf33*, *sDf34* and *sDf35*) uncover all of the region except for the zone between *sDf30* and *sDf35* (see Figure 3). All of the above deficiencies are balanced over *eT1* and are in strains that also carry *dpy-18*, and all but the strain containing *sDf35* carry *unc-46*. Males containing lethals on LGV(left) [*dpy-18/eT1(III);unc-46 let-x/eT1(V)*] were crossed to the above set of

deficiencies. The absence of fertile Dpy Unc (Dpy-18 for *sDf-35*) progeny indicated failure to complement. The lethals were thus roughly mapped. Appropriate other deficiencies were then used, where necessary (see Figure 5 and Figure 6), to identify the zone(s) into which each mutation falls. Once a lethal had been mapped to a zone(s), it was complementation tested against alleles of all known genes in that zone(s) and if necessary to alleles of neighbouring zones: *e.g.* a) if the mutation falls in zone 15, zone 16 has to be tested; b) if the mutation is in zone 17 the gamma irradiation induced mutations (putative deficiencies) in zone 18 have to be tested. The results of the complementation tests allowed for the assignment of the mutations to appropriate genes.

(ix) *Two-factor mapping to unc-46*

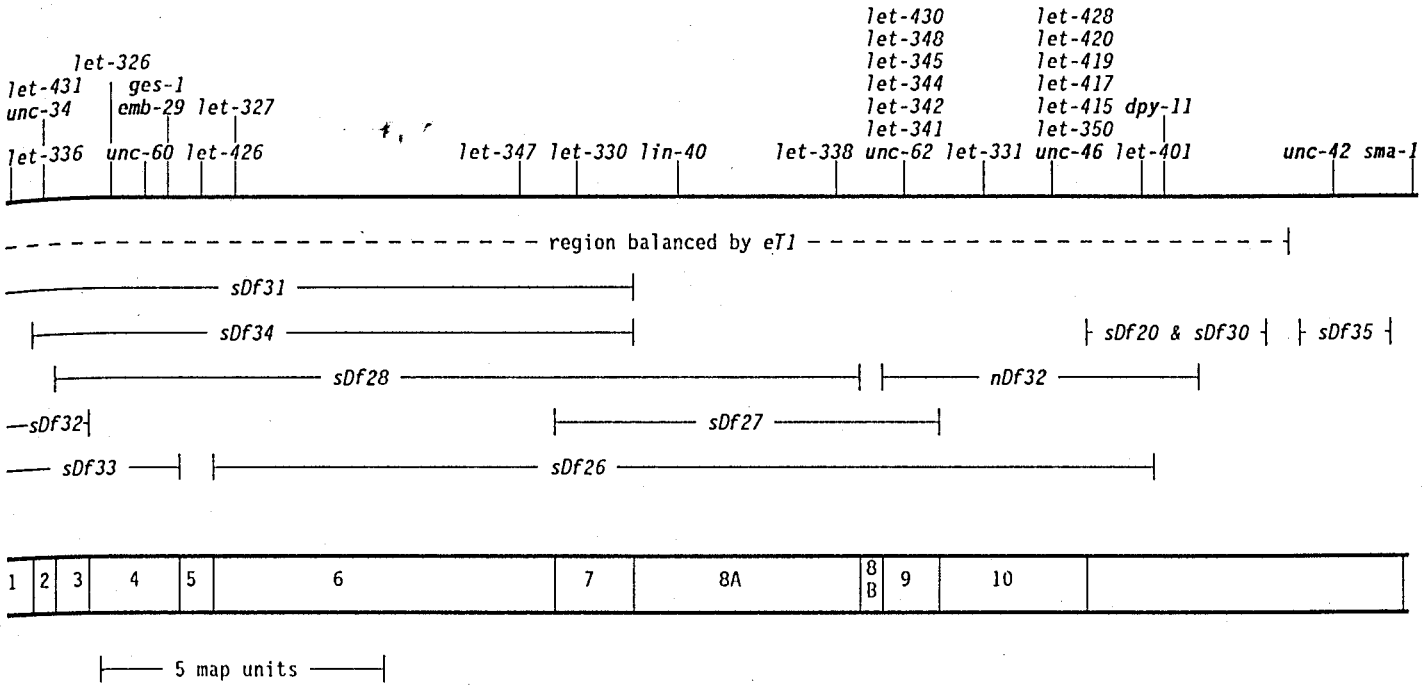
All the mutations are linked to *unc-46(e177)*. To obtain two-factor map distances hermaphrodites heterozygous for *let* mutations over +, were picked individually and brooded every 12 hours until they were exhausted of progeny. Rose and Baillie (1979) showed that recombination is temperature dependent; therefore, all brooding was done at 20C. All mature F_1 progeny were scored. The map distance between the lethal mutations and *unc-46* was calculated as $100\{1-[1-3U/(U+W)]^{1/2}\}$, where U = the number of F_2 Uncs and W = the number of F_2 wild-types.

Figure 3

Genetic map of LGV(left) from Rosenbluth *et al.* (1988). This is the genetic map of LGV(left) prior to the work in this thesis.

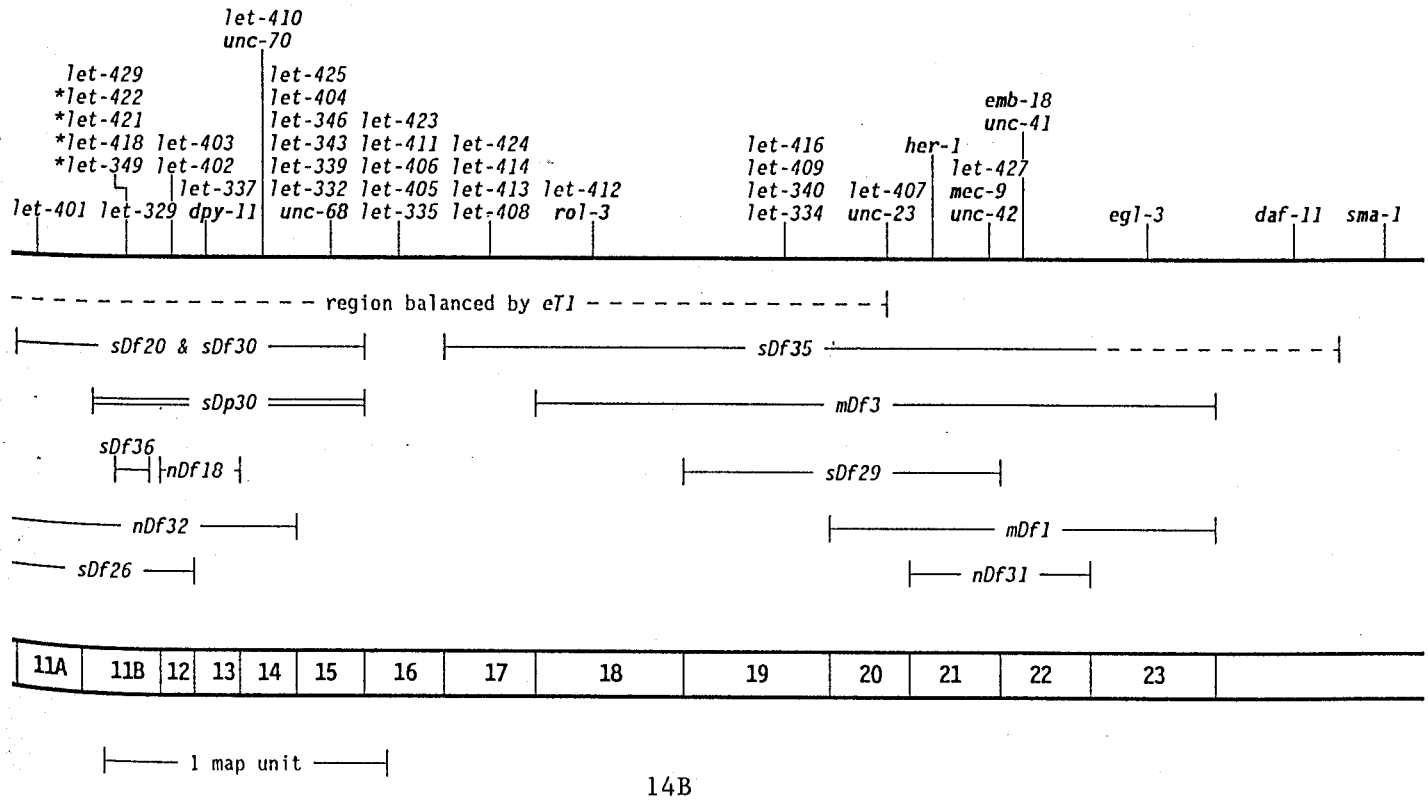
A

Zones 1-10



B

Zones 11-23



Section 1:

Formaldehyde mutagenesis: Dose-response curve and the analysis of mutational events

INTRODUCTION

The main subject of this Thesis is the mapping of the majority of the essential genes on 23 m.u. of the left half of LGV in *C. elegans* that is balanced by the reciprocal translocation *eT1(III,V)* (Rosenbluth and Baillie 1981). A set of deficiencies is essential for this work. A set of large gamma irradiation induced deficiencies was already available. These deficiencies divided LGV(left) into 22 zones (see Figure 3 taken from Rosenbluth *et al.* 1988). A further subdivision of these zones would facilitate the mapping of newly acquired ethyl methanesulfonate (EMS) induced lethal point mutations. Since formaldehyde was demonstrated to produce deficiencies in *D. melanogaster* (Slizynska 1957; O'Donnell *et al.* 1977) and since Moerman and Baillie (1981) have shown that formaldehyde induces small deletions in *C. elegans*, formaldehyde was chosen to induce a new set of small deficiencies in LGV(left). In this section a formaldehyde dose-response curve is developed and a number of formaldehyde induced mutations including some deficiencies are isolated and characterized.

Formaldehyde induced mutagenesis in *C. elegans* can be analyzed at the same time that the deletions are screened for. To do so, the eT1 system (Rosenbluth *et al.* 1983; 1985) can be used. The eT1 system screens for recessive lethal mutations induced in the region balanced by *eT1* which includes at least 20 m.u. of LGIII(right) in addition to the 23 m.u. of LGV(left). The break on LGIII maps to a location very close to (or in) *unc-36* and the homozygous *eT1* individual has an Unc-36 phenotype. *eT1* is stable. It has been intensively studied for several years and has never been observed to break down. It is an excellent system for testing mutagens and has been used to develop dose-response curves for the mutagens EMS, gamma irradiation (Rosenbluth *et al.* 1983), ionizing radiation (Nelson *et al.* 1989) and ultraviolet irradiation (UV) (Stewart *et al.* 1990).

Various concentrations of formaldehyde were used to develop a dose-response curve for the induction of lethals. In doing so, 112 strains containing lesions in or very close to the *eT1* balanced region were isolated and a subset of those mutations were genetically analyzed. These lethals were compared to lethals induced by either 0.012M EMS or 1500R gamma irradiation in a number of respects. The number of lethals that mapped to each of the two linkage groups were compared. The loss of expression of markers in lethal bearing strains was tested (see Results). The distribution of "outcrossed" (defined in Materials and Methods) brood sizes were compared. A number of the simple formaldehyde induced lethals on LGV were positioned and analyzed to discern if they were deficiencies, putative point mutations that define new genes or putative point mutations in known genes. This work has been published (Johnsen and Baillie 1988).

MATERIALS AND METHODS

(i) *Mutagenesis with formaldehyde*

A formaldehyde solution was prepared according to the method of Moerman and Baillie (1981) by warming 5g of paraformaldehyde (Fisher #T353) in 50 ml of 65C distilled water and then adding NaOH to clear the solution. The solution was adjusted to pH 7.2 with HCl and then diluted to 500 ml by adding M9 buffer (Brenner 1974) giving a 1% (wt/vol) solution. This solution was further diluted with M9 buffer to give the appropriate formaldehyde concentration. The concentrations were chosen based on Moerman and Baillie's (1981) findings. They had shown that a 1% concentration of formaldehyde is lethal to *C. elegans*, but concentrations of 0.1% and 0.07% induced both point mutations and deficiencies in the *unc-22(IV)* region. For each mutagenesis run, a fresh 1% formaldehyde solution was made up.

BC2200 nematodes were washed off the stock plates with M9 buffer. The worms were then collected by centrifugation. They were suspended in either a 0.0%, 0.07%, 0.11%, 0.14% or 0.18% solution of formaldehyde and left at room temperature (20-23C) for 4h. The worms were then removed from the solution with a Pasteur pipette and spotted onto a Petri plate. After 2h, young adult wild-type worms were placed on Petri plates (either 1 per plate or 3 per plate) with lawns of OP50. These worms were left for 3 days and their F_1 progeny were screened for lethal mutations as described in General Materials and Methods.

(ii) *Outcrossed brood sizes*

dpy-18(III)/+;unc-46(V)/+ hermaphrodites give a mean of 332 progeny when not mated (see Results). Ten out of sixteen progeny from an *eT1* heterozygous hermaphrodite are aneuploid and do not survive; therefore, the total brood size of a *dpy-18/eT1;unc-46/eT1* line is approximately 125. The lethals are maintained as heterozygotes over *eT1*. When

they were mapped, the translocated *eT1* chromosomes were replaced with "normal" chromosomes thus increasing the brood sizes. The resulting brood sizes are called the "outcrossed brood" sizes. An outcrossed brood size is called small if it is less than 151, medium if it is greater than 150 but less than 291 and large if it is greater than 290 (only 1 or 2 lines were scored per strain). A recessive lethal with no dominant effect should give an outcrossed brood size of 250. I postulate that medium outcrossed brood sizes are indicative of point mutations or small deficiencies and that small outcrossed brood sizes are associated with more complex mutational events or very large deficiencies. The assumptions about the small and medium outcrossed brood sizes are supported by the results of Rosenbluth *et al.* (1985). Their smallest outcrossed brood size for a point mutation or a deficiency was 145 [*sDf29*(V)] with most outcrossed brood sizes being greater than 200. Their largest outcrossed brood size for a translocation or duplication was 166 [*sT3*(III)].

(iii) *Test for the presence of markers*

Several of the strains had lethal mutations which did not separate from their linked marker by recombination. That is, when attempting to two factor map the mutations with respect to the markers, no homozygous marker progeny appeared. In order to ascertain whether the marker was still expressed in strains where the mutations did not separate from their markers, eT1 bearing heterozygous zygotes from these strains were crossed to BC1958 (*dpy-18/eT1;unc-46/eT1*) males. Because aneuploid progeny from heterozygote *eT1* worms die either as eggs or at an early larval stage, normally the only F_1 worms that survive to adulthood have these phenotypes: Wild-type, Unc-36 or Dpy Unc. The presence of F_1 Dpy Unc progeny from the above cross indicated that both markers were still expressed. F_1 Dpy-18s indicated that the *unc-46* marker was not expressed, F_1 Unc-46s indicated that the *dpy-18* marker was not expressed, and the presence of Dpy Uncs,

Unc-46s and Dpy-18s indicated that at least one of the markers was involved in a more complicated mutational event than a point mutation or a deficiency.

Figure 4

Outcrossed brood sizes for 34 0.012M EMS, 85 0.07%, 0.11%, 0.14% and 0.18% formaldehyde, and 85 1500R gamma irradiation induced lethal mutations in the *eT1* balanced region.

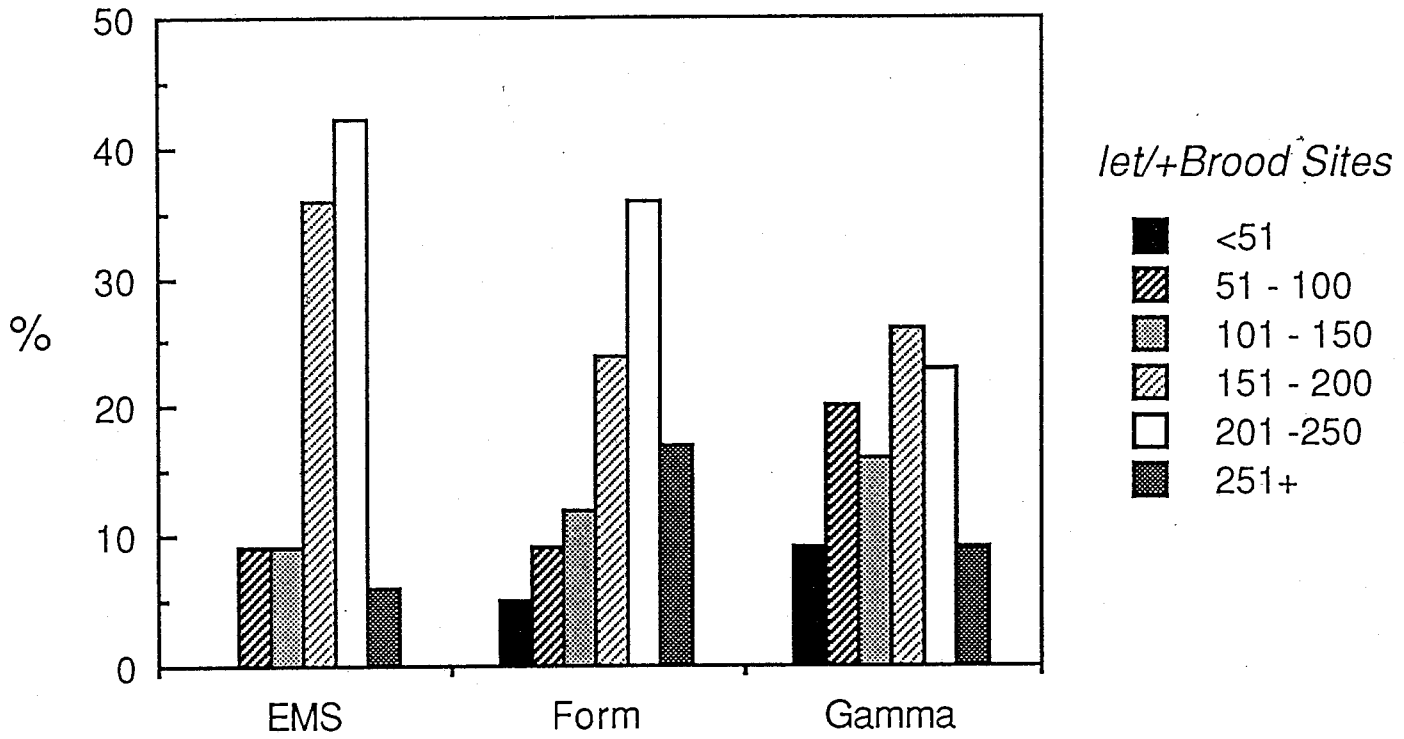

50
40
30
20
10
0
%
EMS
Form
Gamma
let/+Brood Sites
<51
51 - 100
101 - 150
151 - 200
201 -250
251+

Figure 5

Genetic map of the *eT1* balanced region of LGV(left) with the addition of the formaldehyde induced mutations.

LINKAGE GROUP Ⅴ (LEFT)

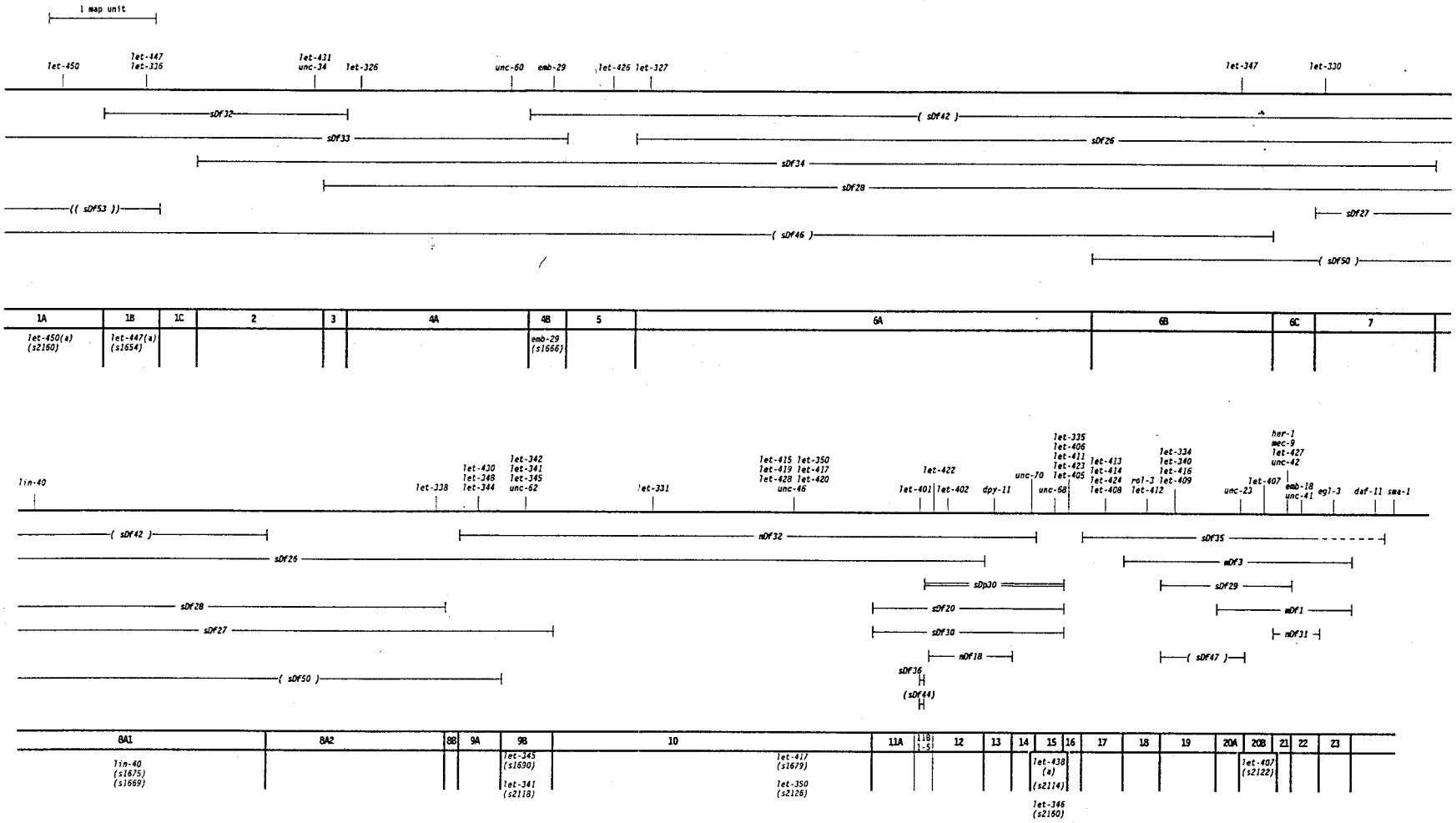
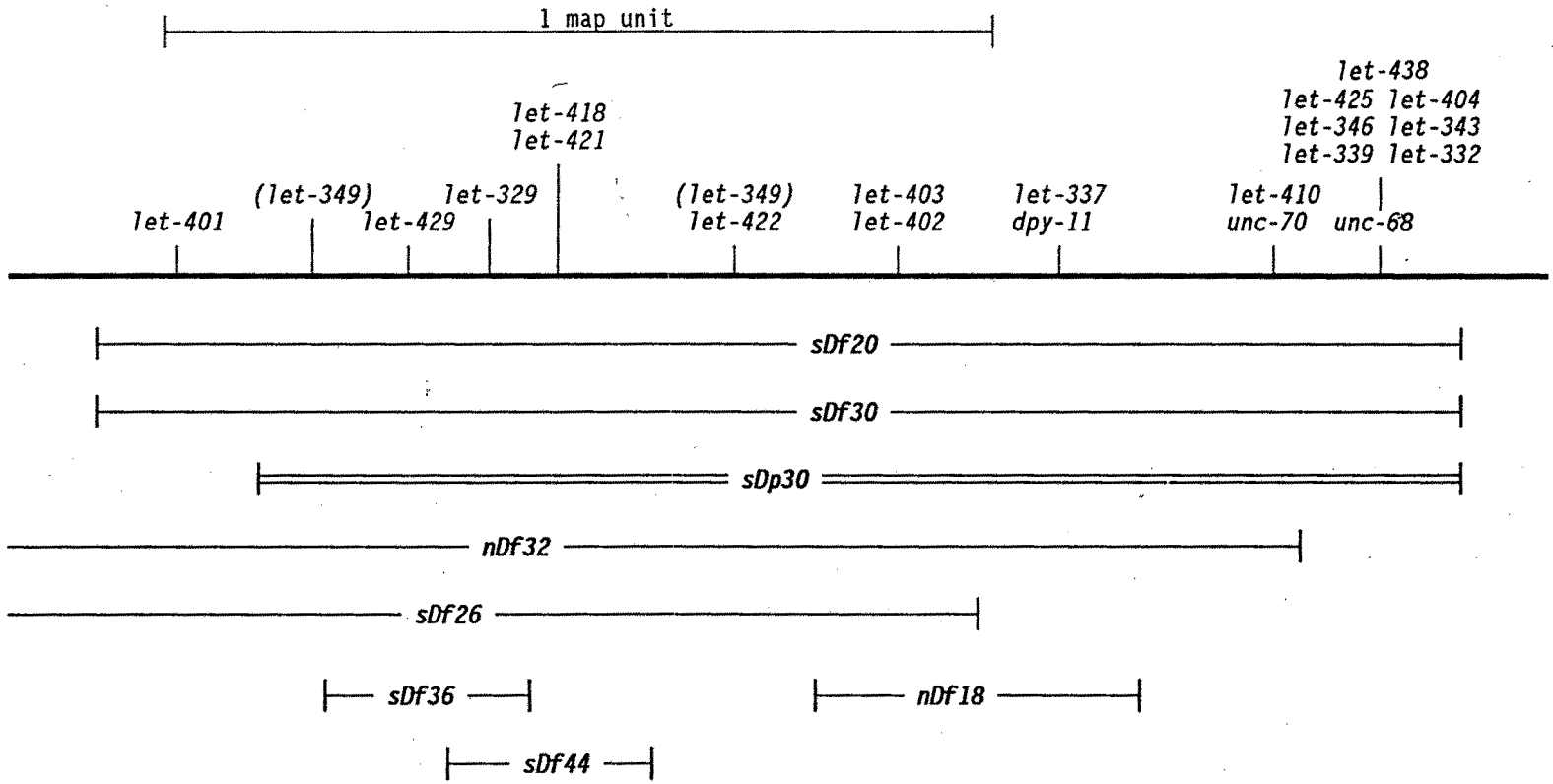
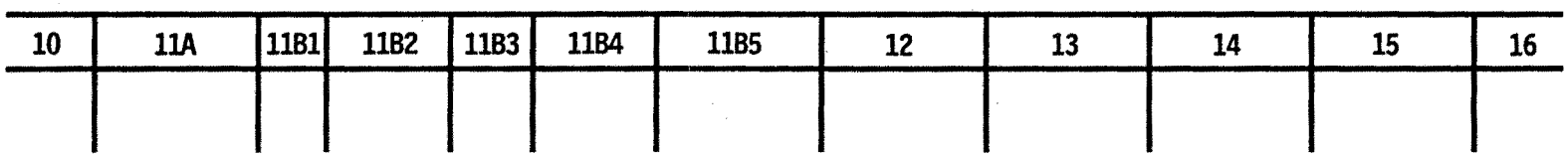


Figure 6

Expansion of zones 11A through 15. *sDf36* and *sDf44* are floating and have not been right/left positioned relative to each other. *let-349* could be in either 11B1 or 11B5, *let-422* has been placed in 11B5 by two factor mapping only (see text). *let-429* could be in 11B5 and if so then *let-418* and *let-421* would be in zone 11B2.

EXPANSION OF ZONES 11A—15





RESULTS

A high incidence of death of P_0s was observed after treating *dpy-18/eT1;unc-46/eT1* worms with formaldehyde, therefore a relatively large number of worms were treated. The number of F_1s from the 0.07%, 0.11% and 0.14% formaldehyde treated P_0s did not vary significantly from the number of the F_1s from non-mutagenized P_0s. The number of F_1 progeny from the P_0 worms treated with 0.18% formaldehyde was approximately 85% of the untreated P_0 brood size. This indicates that physiological damage plays a major role at higher doses.

The percent of sterile F_1 individuals increased linearly with formaldehyde dose and in all cases was less then the 2.5% sterility rate observed for 0.012M EMS treated worms (see Section 4). Therefore F_1 sterility does not pose a significant problem.

The dose-response data are given in Table 1. The "dose" refers to the concentration of formaldehyde that the worms were exposed to for 4h at room temperature. The "response" is the percentage of fertile F_1s that carried at least one lethal mutation within the region balanced by *eT1*. The lethal induction rate for 0.11% formaldehyde is approximately 1.6% (see Table 1). This is similar to the rate for 0.004M EMS (26/1747 or 1.5%) and greater than the rate for 500R gamma irradiation (46/4789 or 1.0%) (Rosenbluth *et al.* 1983).

To establish a value for the outcrossed brood size of *dpy-18/+;unc-46/+* hermaphrodites, all the adult progeny of seven individuals were scored. The mean brood size was 332. The wild-type:Dpy-18:Unc-46:Dpy Unc ratio that was well within the 95% confidence limits for a 9:3:3:1 ratio was obtained. J. Hodgkin (1988) scored a mean self-progeny brood size of 329 for N2 hermaphrodites).

This shows that if either marker has a detrimental effect, the effect is very small. There is no significant difference in the percent of lethals giving small (<151) outcrossed

brood sizes between the low (0.07% and 0.11%) and high (0.14% and 0.18%) formaldehyde doses.

In Figure 4 the outcrossed brood sizes for lethals induced by three different mutagens are compared. The lethals in Figure 4 do not include those that block development at late larval or adult sterile stages. The mutagens used were: 1500R gamma irradiation (Rosenbluth *et al.* 1985 and L. M. Addison personal communication), 0.012M EMS (Section 4) and formaldehyde. EMS generated lethals produce mainly medium outcrossed brood sizes: 82% have between 151 and 290 offspring. Gamma irradiation produced lethal bearing heterozygotes which varied greatly in outcrossed brood size, with only 56% between 151 and 290. The formaldehyde induced lethal bearing heterozygotes gave a distribution of outcrossed brood sizes which fell between that of EMS and gamma irradiation, with 76% having medium size broods.

The formaldehyde induced lethal mutations of 96 out of a total of 112 strains were mapped to linkage group. 34% lie on LGV(left) and 66% lie on LGIII(right). This includes the strains that have mutations that map to both LGIII and LGV but does not include second site hits on one chromosome. A similar ratio was found for 0.012M EMS induced lethals: 35% on LGV(left) and 65% on LGIII(right) from a sample of 147 lethal bearing strains Johnsen *et al.* (1986). A larger sample of 751 EMS induced lethals gives the same fraction mapping to each chromosome (see Section 4). Two 1500R gamma irradiation mutagenesis screens (Rosenbluth *et al.* 1985; L. M. Addison personal communication) produced 67 lethal bearing strains with lesions restricted to one chromosome, 40% of the lethals lie on LGIII(right) and 60% lie on LGV(left).

When the formaldehyde induced lethals were mapped to either LGIII or LGV, 18 strains were identified that produced either no Dpy-18 or no Unc-46 progeny. Such strains could arise in two possible ways. 1) The lethal is tightly linked to the marker or 2) the marker can not be expressed in the strain. The strains were tested for the presence of

markers (see Materials and Methods). Fifteen strains proved to still express both markers, indicating a lethal tightly linked to a marker. Two strains failed to express one of the markers: BC2772 did not express *dpy-18* and BC2482 did not express *unc-46*. This suggests that the wild-type allele of *dpy-18* in BC2772 and the wild-type allele of *unc-46* in BC2482 may each have been translocated from its original location(s) on the *eT1* chromosomes to a location that is tightly linked to the suppressed marker, or a formaldehyde induced crossover had occurred between the region containing the marker and the region containing its wild-type allele. One other strain (BC2475) produced Dpy-18, Unc-46 and Dpy Unc progeny. It is possible that either *dpy-18* or *unc-46* has translocated elsewhere in the genome and has not been lost. These three strains demonstrate that formaldehyde can cause complex mutational events.

Thirty-one of the 96 strains containing formaldehyde induced mutations that mapped in or very near the *eT1* balanced region have lesions that are strictly on LGV(left). Five strains were lost. Four of these had low brood sizes, were very sickly, and died out. The fifth was outside the balanced region and crossed away. Three of the remaining strains have markers that behave irregularly (BC2772, BC2482 and BC2475 discussed above). Three more are sickly and have small brood sizes. They probably contain complex mutational events and therefore have not been analyzed futher. One strain has a medium brood size, is fertile but develops very slowly and therefore analysis was stopped. Initially the latter strain had a lethal mutation, but it was probably outside the *eT1* balanced region and crossed away revealing the mutation that causes slow development. Four strains have lethal mutations that span or are just outside the *eT1* balancer boundary. The analysis of three of these is incomplete. This leaves fifteen strains. The lethals in five map as deficiencies. Ten strains contain eleven putative point mutations (one strain has two putative point mutations), three of which define new genes and eight of which are alleles of known genes (see Table 2).

All five of the deficiencies and seven of the ten strains containing putative point mutations have medium outcrossed brood sizes. The other three strains containing putative point mutations have late larval lethal or sterile adult phenotypes and their outcrossed brood sizes were not scored. The two strains that failed to express a marker (BC2772 and BC2482) both have small outcrossed brood sizes whereas BC2475 has a medium outcrossed brood size.

The map of LGV(left) (Figure 5 and Figure 6, for a more complete map see Figure 12) is based on the map in Figure 3 of Rosenbluth *et al.* (1988). The chromosome, with genes positioned on it, is at the top of the map. The chromosomal rearrangements that break the chromosome into zones are in the middle and the zone designations are on the bottom of the map. The map distances of the genes in zone 1 (*let-450*, *let-447* and *let-336*) from each other and from *unc-34* are not known. The distances shown for these genes are arbitrary. The formaldehyde induced deficiency names are shown on the map in single parentheses *e.g.* "(*sDf42*)". Lesions that previously were putative point mutations and have now been shown to be deficiencies by formaldehyde induced mutations are on the map with their names in double parentheses *e.g.* "((*sDf53*))". Formaldehyde lesions that are putative point mutations are listed at the bottom of the map in the appropriate zone with their allele names (in parentheses) following their gene name. Formaldehyde lesions that define new genes have the letter "a" in parentheses following the gene name *e.g.* "*let-450*(a) (*s2160*)". One of the deficiencies (*sDf31*) shown on the Rosenbluth *et al.* (1988) map is not being used. *sDf31* was isolated from a strain with an N2/BO background (Donati 1985; Rosenbluth *et al.* 1988). It may still have active Tc1 transposons because it occasionally gives progeny with unexpected phenotypes. Because it was desirable to use stable and predictable strains to define the zones of LGV the strain containing *sDf31* was not used.

The five formaldehyde induced deficiencies are designated *sDf42*, *sDf44*, *sDf46*, *sDf47* and *sDf50*. *sDf46* uncovers the left-most known gene on LGV (*let-450*) suggesting that *sDf46* deletes the end of the chromosome. *sDf46* breaks between *let-347* and *sDf27* and between zone 6B and 7.

sDf42 breaks between *unc-60* and *emb-29* splitting zone 4 into 4A and 4B. The right end of *sDf42* breaks between *lin-40* and *let-338* creating zones 8A1 and 8A2 (see Figure 5).

sDf50 breaks between *let-327* and *let-347* dividing zone 6 into 6A and 6B. The right break point of *sDf50* divides zone 9 into 9A and 9B. *let-344*, *let348* and *let-430* are in 9A and therefore to the left of *let-341*, *let-342*, *let-345* and *unc-62* which are in 9B. The relative positions of genes in zone 11B are not yet clear. *sDf44* is a short deficiency that lies entirely in zone 11B. It partially overlaps *sDf36* which also lies entirely in 11B creating zone 11B3. *let-329* lies in the overlap region (11B3). *let-418* and *let-421* are uncovered by *sDf44* but not by *sDf36*. It is not yet know if *sDf44* overlaps with the right end of *sDf36* as shown on the map (Figure 6) or the left end of *sDf36*. Therefore, both *let-418* and *let-421* may lie to the left of *sDf36* and *let-429* may lie to the right of *sDf44*. *let-422* was placed in zone 11B5 because 2-factor mapping from *dpy-11* gave a distance of 0.3 m.u. (0.1-0.9) (Rosenbluth *et al.* 1988) but *let-422* could be in zone 11B1. *let-349* could be in either zone 11B1 or 11B5 and so was placed it in both zones on the map (Figure 6).

The left end of *sDf47* has not been separated from the left end of *sDf29*. However, on the right end, *sDf47* complements *nDf31*, splitting zone 20 into zones 20A and 20B. Zone 20A contains *unc-23* and 20B contains *let-407* (see Figure 5).

The formaldehyde screens produced 10 strains containing 11 putative point mutations on LGV. One strain, BC2821, carried a late larval lethal, *s2160*. Upon complementation testing BC2821 against the set of deficiencies it was found to contain two mutations, one of which is allelic to *let-346* in zone 15 and has been assigned the allele

name *s2166*. The other lesion retained the *s2160* designation. Initial testing (in 1987) showed that *s2166* failed to complement *sDf33*, *sDf53* and several other deficiencies but it complements for *sDf34* and *sDf32*, thus showing that *sDf33* extends to the left of *sDf32*. Subsequent retesting (in 1989) showed that *s2160* complemented *sDf33* and *sDf53*. Therefore *s2160* has been lost possibly through reversion. *s2160* defined a new essential gene named *let-450* which lies in zone 1A. *let-450* and zone 1A are the left most known gene and zone respectively on LGV. *s2160* complements one allele of *let-336* but fails to complement another (*s957*). *s957* has therefore been designated as a deficiency (*sDf53*) that splits zone 1B into 1B and 1C.

s1654 defines a new gene, *let-447*, in zone 1B. *s1654* behaves like a hypermorph. When it is homozygous it usually causes sterility although sometimes a few eggs that do not hatch are laid, but over a deficiency, *s1654* containing strains consistently give many more eggs although none of them hatch. This effect was observed with *sDf32*, *sDf33*, *sDf46*, *sDf53* and several other deficiencies that uncover the same region. Two possible explanations for this are discussed in Section 4.

s1666 is an allele of *emb-29* in zone 4B. Two alleles of *lin-40* (zone 8A1) were isolated; they are *s1669* and *s1675*. *s1669* blocks as a late larval lethal and *s1675* blocks as a sterile adult but occasionally a few nonviable eggs are laid. *s1675* was mated to wild-type (N2) males to test for male rescuability. The matings were unsuccessful, possibly due to vulva abnormalities.

s1690 is allelic to *let-345*. *s2118* is allelic to *let-341*. Both *let-345* and *let-341* are in zone 9B. *s1679* is allelic to *let-417*. *s2126* is allelic to *let-350*. Both *let-417* and *let-350* are in zone 10. *s2114* maps to zone 15 but complements all known genes in that zone, therefore *s2114* was assigned a gene name: *let-438*.

Table 1

(Induction of recessive lethals at various formaldehyde doses)

Dose	Date of Expt.	Number of F_1s tested	Number of lethals	Lethals (%)
Control	21/06/81	1812	2	0.11
	27/09/82[a]	1386	0	0.00
	16/03/87	641	0	0.00
		3839	2	0.05
Formaldehyde				
0.07%	12/05/86	3670	10	0.27
	13/08/86	1163	4	0.34
		4833	14	0.29
0.11%	22/06/86	3682	58	1.58
0.14%	4/09/86	992	16	1.61
0.18%	4/09/86	530	13	2.45
	11/16/86	956	11	1.15
		1486	24	1.62

[a] Rosenbluth *et al.* (1983)

Table 2

(Distribution and type of formaldehyde induced recessive lethal mutations)

Dose (%)	Number Screened	Number of Mutations recovered	Number Mapped to LGV	Deficiencies	Alleles of known genes	Alleles of new genes	Putative Complex mutations	Not Done
0.07	4833	14	2		1		1	
0.11	3682	58	24	5	5	2	9	3
0.14	992	16	3		1			2
0.18	1486	24	2		1	1		
Total	10993	112	31	5	8	3	10	5

DISCUSSION

Formaldehyde was shown to be a mutagen in the 1940s (Rapoport 1946; Kaplan 1948) but its mechanisms of action were not understood and are still not completely clear (Auerbach *et al.* 1977). It is known that formaldehyde reacts with amino groups in denatured regions of DNA. There is another slower reaction that results in the formation of methylene cross-links between amino groups in DNA (Feldman 1975). Formaldehyde can also form reactive hydroxyalkyl peroxides and/or free radicals in the presence of oxidizing molecules. Auerbach *et al.* (1977) proposed that formaldehyde may produce its effects by various mechanisms depending on the organism and mode of application.

Auerbach *et al.* (1977) argued that the most likely reaction with formaldehyde in biological systems is a condensation reaction with amino groups. This suggests that proteins act as a sink for formaldehyde but the organism can only absorb a certain amount of formaldehyde before physiological damage affects it and its germ cells. The observation that the F_1 brood size decreased at the higher formaldehyde dose (0.18%) supports this proposal.

Treatment with formaldehyde has been reported to produce several types of mutations in *D. melanogaster* (Slizynska 1957). The author cytologically analyzed nuclei from the salivary glands of female larvae carrying chromosomes from formaldehyde treated parents. She identified translocations, inversions, deficiencies, repeats, duplications and cases of chromosome loss. She also found that the majority of her formaldehyde induced rearrangements were intrachromosomal (82%). Of 114 mutations 25% were deficiencies and 39% were duplications or repeats. Slizynska also noted that formaldehyde induced approximately 15 times more deficiencies than X-rays at doses that induced similar proportions of sex-linked lethals. O'Donnell *et al.* (1977) confirmed Slizynska's results. They reported 67% (12 out of 18) formaldehyde mutations in the *D.*

melanogaster Adh region were deficiencies. Moerman and Baillie (1981) used formaldehyde to induce *unc-22(IV)* mutations in *C. elegans*. They found that at least 35% (14 out of 40) of those mutations were deficiencies. The remaining 26 were classified as point mutations. Three of those putative point mutations have since been shown to be internal duplications in *unc-22* (T. Starr pers. comm.). T. Starr showed these duplications to be genetically unstable, they revert at a frequency of 10^{-5}. None of my formaldehyde induced mutations are unstable, therefore I probably have not isolated any duplications. From Figure 12 it can be observed that formaldehyde induced deficiencies delete on average about half as many genes as do 1500R gamma irradiation induced deficiencies. Therefore formaldehyde has less drastic mutagenic effects than 1500R gamma irradiation and coupled with the fact that formaldehyde is much cheaper to use than gamma irradiation, it would appear that formaldehyde is an excellent alternative mutagen for the induction of deficiencies.

A dose-response curve for a mutagen allows one to derive the optimum dose. From the formaldehyde dose-response curve, I recommend a formaldehyde dose of 0.1% for the generation of deficiencies. Table 1 shows the lethal mutation rate at each of the tested concentrations of formaldehyde. These results are similar to those of Kaplan's (1948), in which he transferred Drosophila larvae to a medium containing formaldehyde. In four of Kaplan's experiments in which the concentration varied two and one half fold there was no dependence on concentration of the lethal induction rate and the rate was 28 times higher than in the control. Unfortunately these experiments depend on Drosophila eating food containing formaldehyde. With *C. elegans*, the worms are placed directly into a buffer solution containing formaldehyde. Therefore, I feel that the conditions of treatment are more readily reproducible.

All the formaldehyde induced deficiencies and all of the formaldehyde induced point mutations counted (7 out of 10) gave medium outcrossed brood sizes, with the putative

point mutations having an average of 45 more progeny than the deficiencies. Two of the three strains containing putatively complex mutational events had small outcrossed brood sizes. It may be possible to use the outcrossed brood size as a crude indicator for the type of mutational event induced.

Two-thirds of the lethal mutations induced by EMS (Johnsen *et al.* 1986 and Section 4 of this thesis) and formaldehyde mapped to LGIII(right), and one-third to LGV(left). This suggests that there are twice as many genes in the LGIII(right) region than in the LGV(left) region even though both regions are approximately the same length recombinationally, each being about 20 m.u. long. With 1500R gamma irradiation the ratio is reversed: 40% of the mutations map to LGIII and 60% map to LGV (Rosenbluth *et al.* 1985; L. M. Addison personal communication). A partial explanation for the discrepancy between the gamma irradiation results and the EMS and formaldehyde results, is that *eT1* has a defect in the *unc-36* gene (Rosenbluth and Baillie 1981) and any lesion uncovering *unc-36* would give an Unc-36 phenotype in an *eT1* hetrozygote and therefore would be ignored when the F_1 wild-types are picked during screening.

However, undetected *unc-36* deletions cannot entirely explain the observed discrepancy for the following reason: from the 67 gamma irradiation induced lethals mapped to chromosome, 26 mapped to LGIII and 41 mapped to LGV. If the same ratio that applies to the EMS and formaldehyde induction of lethals also applies to the induction of lethals by gamma irradiation, then one would expect two-thirds of these lethals to map to LGIII and one-third map to LGV. For 41 lesions on LGV one would expect about 82 on LGIII. There are only 26. Therefore there are approximately 56 too few lesions. If all the "missing" lesions were undetected *unc-36* deletions then 68% (56/82) of all gamma induced lesions on LGIII(right) would delete *unc-36*. This is far too many and therefore I propose the following hypotheses to account for the "missing" lesions. One hypothesis requires differential sensitivity of chromosomal areas to gamma irradiation whereas the

other two do not. 1) the *unc-36* region is very sensitive to gamma irradiation, resulting in a disproportionately large number of lesions that affect *unc-36*. These would not be detected in our screen; 2) there is a haplo-lethal area on LGIII and any mutations that remove it are inviable as heterozygotes and would not be picked up in our screen. Mains *et al.* (1990) suggest that *ct45*(III) may identify a haplo-insufficient gene in the LGIII(right) region. 3) the LGV(left) balanced region is composed of more DNA than LGIII(right) but has a lower gene density.

Five deficiencies and 11 putative point mutations induced by formaldehyde have been isolated. Only 3 of the 11 putative point mutations define new genes. A small set of EMS induced lethals were analyzed in parallel with the formaldehyde induced lethals. 60% of the EMS induced lethals defined new genes, whereas less than 30% of the formaldehyde induced lethals defined new genes. Although this difference is not statistically significant it does suggest that some of the formaldehyde induced putative point mutations may be deficiencies that delete a small number of genes.

The formaldehyde induced mutations have increased the number of zones in the *eT1* balanced region of LGV (1-20 inclusive) from 22 (Rosenbluth *et al.* 1988) to 34. The optimum number of zones for the analysis of this region should be 40-50. Several more putative deficiencies have been analyzed and are reported on later in this Thesis.

When looking at the recombination map of LGV(left) it can be seen that the deficiencies in the gene cluster (*dpy-11-unc-42*)are much shorter than the deficiencies to the left of the cluster. This may reflect a much lower number of DNA base pairs per m.u. near the end of the chromosome than in the middle of the chromosome.

Section 2:

Analysis of Lethal mutations induced

by the Transposon Tc1

INTRODUCTION

The ability to induce mutations that molecularly tag genes is invaluable both for cloning genes and for correlating the genetic and molecular maps. In *C. elegans* the transposon Tc1 has been used to molecularly tag specific genes (see Greenwald 1985; Moerman *et al.* 1986 for examples; for a review on Tc1 see Moerman and Waterston 1989). Tc1 insertion may be non-random because Tc1 has a "strong" nine base pair consensus sequence (Mori *et al.* 1988a; Eide and Anderson 1988) and it also inserts into some genes at relatively high frequency [examples include *unc-22* (Moerman *et al.* 1988) and *unc-54* (Eide and Anderson 1985)] which both show a high frequency of spontaneous Tc1 induced mutations in the *C. elegans* BO (var. Bergerac) strain. An analysis of Tc1 induced mutational events in a large region of the genome could provide a spectrum of the types of lethal mutations induced by Tc1 and could provide a set of anchor points useful for correlating the genetic and molecular maps.

There are approximately 300 copies of Tc1 in BO and only about 30 in the wild-type N2 (var. Bristol) strain (Emmons *et al.* 1983). Tc1 mobilizes readily in BO but not in N2 (Moerman and Waterston 1984; Mori *et al.* 1988b; Collins *et al.* 1987). Loci called mutators (*mut*) are responsible for the mobilization of Tc1 and some have been localized in BO (Mori *et al.* 1988b). A mutator strain (RW7037) with about 70 copies of Tc1 and about one half the Tc1 mobility of BO was derived from BO by crossing to the N2 strain. The mutator in RW7037 is on LGI and called *mut-4(st700)* (Mori *et al.* 1988b). Tc1 induced mutations have been recovered from RW7000, from which RW7037 was derived (Moerman *et al.* 1988).

A screen for isolating Tc1 recessive lethal mutations in the *nT1* balanced regions of LGIV(right) and LGV(left) using *mut-4(st700)* was undertaken by Clark *et al.* (1990). *nT1* is a translocation that suppresses recombination over a total of approximately 49 map

units (m.u.) on LGIV(right) and LGV(left) (Ferguson and Horvitz 1985; Clark *et al.* 1988). The *nT1* recombinationally suppressed region on LGV(left) includes the entire region that is balanced by *eT1* (Clark 1990).

Thirty-seven independently derived lethal mutations had been recovered after screening 3503 F_1 "chromosomes". Twenty-eight of the lethal mutations had mapped to LGIV or LGV (Clark 1990). Only four (14%) mapped to LGIV, while 24 (86%) mapped to LGV. This contrasts to EMS induced recessive lethal mutations where 57% are on LGIV and 43% on LGV (Clark 1990).

In this Section the analysis of the Tc1 induced lethal mutations on LGV(left) is discussed.

MATERIALS AND METHODS

(i) *Genotype used for mapping lethals on LGV*

After removing *mut-4*, Clark (1990) mapped 24 Tc1 induced lethal mutations to the *nT1* balanced region of LGV(left). The genotypes of the strains containing these mutations was *unc-22/nT1(IV);(let-x)unc-46/nT1(V)*. In order to map these mutations on LGV, male strains of the genotype *dpy-18/eT1(III);(let-x)unc-46/eT1(V)* were constructed.

(ii) *Mapping to location on LGV(left)*

LGV(left) has been divided into zones by a set of rearrangement breakpoints (Rosenbluth *et al.* 1988; Johnsen and Baillie 1988; Clark *et al.* 1990). A subset of these deficiencies (*sDf26, sDf30, sDf33, sDf34* and *sDf35*) uncovers all of the region except for the region between *sDf30* and *sDf35* (zone 16 see Figure 12). All of the above deficiencies are balanced over *eT1* and are in strains that also carry *dpy-18*, and all but the strain containing *sDf35* carry *unc-46*. Males from the sample of Tc1 induced lethals on LGV(left) were crossed to hermaphrodites of the above set of deficiencies. The absence of fertile Dpy Unc (Dpy-18 for *sDf35*) progeny indicated failure to complement. The lethals were thus roughly mapped. Other appropriate deficiencies were then used, where necessary (see Figure 12), to identify the zone(s) into which each mutation falls. Once a lethal had been mapped to a zone, it was complementation tested against alleles of all known genes in that zone. The results of the complementation tests allowed for the assignment of the mutations to appropriate genes.

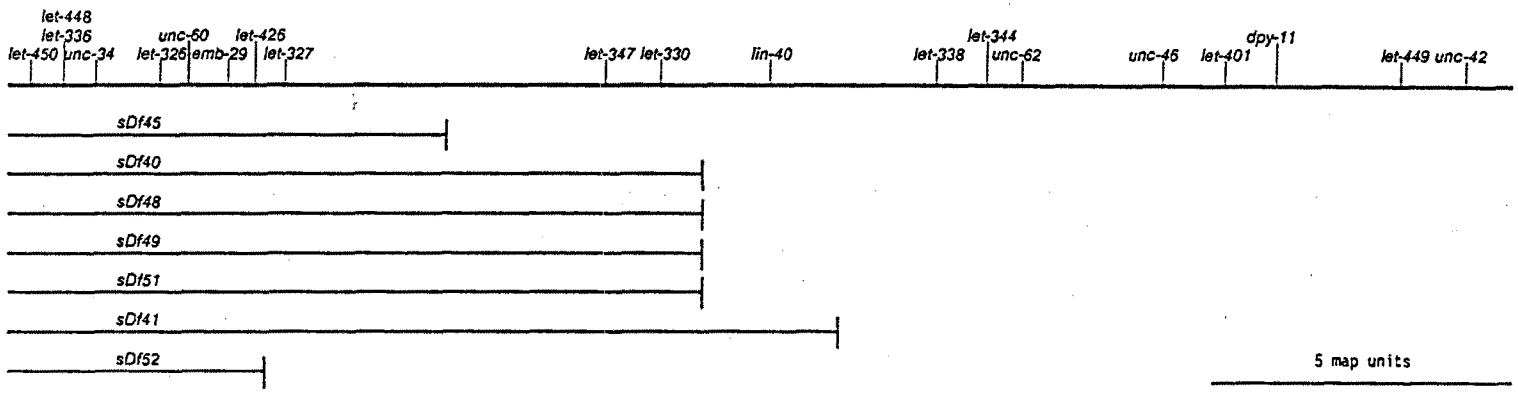
(iii) *Association between deficiencies and half-translocations*

m denotes a visible marker(s) on the chromosome being tested: *unc-35 or dpy-5* on LGI; *unc-52* on LGII; *unc-64* on LGIII; and *unc-26 or dpy-4* on LGIV (See Figure 1). *m/+; Df unc-46/+ + (V)* hermaphrodites were brooded every 24h and all progeny scored. *Dfs* recombine away from *unc-46* at very low frequency (see Section 3), therefore the presence

of Unc-46 individuals does not interfere with the scoring. One quarter of the resultant progeny being of M phenotype indicated failure to complement *m* by the deficiency bearing chromosome. One twelfth M phenotype indicated complementation and therefore the deficiency was associated with a translocation of the chromosome containing *m*.

Figure 7

Partial genetic map of the left end of LGV showing the set of *mut-4* induced deficiencies and two new *mut-4* induced complementation groups (*let-448* and *let-449*).


let-448
let-336
let-450
unc-34
unc-60
let-426
let-325
emb-29
let-327
let-347
let-330
lin-40
let-344
let-338
unc-62
unc-46
dpy-11
let-401
let-449
unc-42
sDf45
sDf40
sDf48
sDf49
sDf51
sDf41
sDf52
5 map units

41B

RESULTS

Upon balancing the 24 LGV mutations over *eT1*, it was found that two mutations were not in the *eT1* balanced region. The remaining 22 mutations were analyzed. It is important to note that these mutations are of independent origin (Clark *et al.* 1990).

Two mutations define new genes; *let-448* and *let-449*. *let-448* maps to zone 1B2 (Figure 12). *let-449* complements all the deficiences in the set noted in Materials and Methods but maps 4.5 (2.1-9.1) m.u. from *unc-46* (Clark 1990); therefore, it was placed in zone 16 (Figure 12). Six mutations are alleles of *lin-40* in zone 8A1 (Figure 12). Thus it appears that *lin-40* is a "hot spot" for Tc1 mutagenesis.

Seven mutations are deficiences (*sDf*s *40, 41, 45, 48, 49 51* and *52*) which break in various places on LGV(left) (see Figure 7). However, they all uncover *let-450* which is the left-most known gene on LGV (see Section 1). Four of these deficiencies have right breakpoints that lie between both *let-330* and *let-461* (not yet right/left positioned relative to each other) and *lin-40* (see Figure 7). These four breakpoints have not been further resolved. *let-330* and *let-461* are currently the nearest known genes to the left of *lin-40*. The deficiency *sDf41* either breaks within *lin-40* or between *lin-40* and the next known gene to the right [either *let-338*, *let-455* or *let-466* (Fig)]. Therefore, not only is *lin-40* a hotspot for *mut-4* induced mutagenesis, but the region around *lin-40* is also very susceptible to *mut-4* induced chromosome breakage. This result suggests that the region around *lin-40* is either very susceptible to Tc1 insertion or that there may be a resident Tc1 element(s) responsible for the mutations in and around *lin-40* (for more on this see Section 4).

It is possible that Tc1 mobilization can cause reciprocal translocations due to recombination between Tc1 elements on different chromosomes. The loss of the chromosome containing the translocated LGV(left) region would result in a

half-translocation. These half-translocations should appear as deficiencies for LGV genes. To find out if the Tc1 induced deficiencies really are associated with half-translocations, autosomal genes with visible phenotypes were chosen to use in complementation tests against *sDf40* and *sDf41* bearing chromosomes. Previous workers have implicated one half of each chromosome as having loci necessary for homolog recognition which is prerequisite for recombination and disjunction (discussed by McKim *et al.* 1988; Rose and McKim 1989; Herman and Kari 1989). Therefore, only translocations not bearing these loci could survive. The genes chosen to test for half-translocations are from the appropriate ends of the other autosomes: *unc-35, dpy-5* on LGI; *unc-52* on LGII; *unc-64* on LGIII; and *unc-26, dpy-4* on LGIV (see Figure 1). *m/+; Df unc-46/++* strains were constructed where *m* represents the genes used for testing. One quarter M progeny from the test strains indicated failure to complement *m* by the deficiency bearing chromosome; one twelfth M indicated complementation and that the deficiency was associated with a half-translocation of a region of the chromosome containing the wild-type allele of *m*. In all cases the homozygous visible mutations were not complemented by the deficiency bearing chromosomes. This strongly suggests that the deficiencies are not associated with half-translocations involving the autosomes.

The final seven *mut-4* induced mutations caused low fertility and were difficult to work with. They were probably caused by complex mutational events, or are large deficiencies, or Tc1 elements were still mobile in these strains. A *mut-4* strain has been isolated that contains additional mutator activity linked to LGIV (Mori *et al.* 1988b). Therefore, some strains carrying *mut-4* induced lethal mutations could also be carrying mutator activity, in spite of the fact that *mut-4* was removed from the strains (Clark 1990).

DISCUSSION

Fifteen Tc1 induced mutations were mapped to location on LGV(left). Two identify new genes (*let-448* and *let-449*), six are alleles of *lin-40* (for more on *lin-40* see Section 5), and seven are deficiencies. The three genes with Tc1 induced mutations are possibly clonable and are potential anchor points that can be used for correlating the genetic and molecular maps of LGV.

There is a "hotspot" for Tc1 induced putative point mutations in *lin-40* and a "hot region" for Tc1 induced deficiencies around *lin-40*. All seven deficiencies are putative tip deletions because they delete the left-most known gene on the chromosome but they are probably not associated with half-translocations. The right breakpoints of the deficiences are clustered in the region around *lin-40*. In contrast, gamma irradiation (Rosenbluth *et al.* 1988), UV (Stewart *et al.* 1990) and formaldehyde (Section 1) mutagenesis result in both tip and internal deficiencies with breakpoints that appear to be randomly distributed.

Mutagenesis in the *zeste-white* region of *Drosophila* using dysgenic P-M hybrids has shown that P element mobility can result in insertions, deficiencies and inversions (Simmons and Lim 1980). The authors also noted that hybrid dysgenesis-induced mutations occurred at preferred sites. Simmons *et al.* (1984) described two major classes of P element mutation mechanisms: 1) local activity of P elements resulting in mutations at or near the site where the P element resides and 2) P element transposition and insertion at a new site in the genome. One explanation for mutational hotspots of P elements comes from studies on a wild-type revertant of a P element induced *yellow* allele (Geyer *et al.* 1988). A residual P element sequence in the revertant allele appears to be a preferred site for the introduction of new P elements, not through insertion but through either recombination or gene conversion. If there are Tc1 sequences within or near *lin-40*, then a similar process may explain the mutational hotspot. Another possibility is that

there is a Tc1 element resident near *lin-40* and its imprecise excision causes the induction of the *lin-40* alleles. This is not very probable because at least three of the six *lin-40* Tc1 alleles are associated with a new Tc1 insertion (section 5).

An alternative hypothesis is that the putative tip deficiencies generally arose through the excision of resident Tc1 elements and that the *lin-40*, and other putative non-deficiency mutations, generally arose through Tc1 insertions. The tip deficiencies could be explained by the mobilization of resident Tc1 elements, if the excision is not followed by repair of the chromosomal breaks. Engels *et al.* (1990) proposed a homolog-dependent model of P element loss in Drosophila. They noted that precise P element loss occurred at a rate hundreds of times greater when the p element was heterozygous as opposed to when it was homozygous. In the model they proposed that P transposition created a double strand break which was repaired by using the homolog chromosome as a template. A similar mechanism could explain the high incidence of tip deficiencies in the *eT1* translocation region of *C. elegans* where the homolog is not available to effect repair. The mobilization of Tc1 induced mutations were picked up over a translocation (*nT1*). Translocations are thought not to pair in *C. elegans* (Rosenbluth and Baillie 1981). Therefore the repair mechanism proposed in the Engel *et al.* (1990) model would not be available and thus any Tc1 excision would be accompanied by a tip deficiency. A method to test this hypothesis would be to isolate mobilization of Tc1 induced recessive lethal mutations in a chromosome region both in the presence and absence of a translocation. A substantially higher ratio of putative tip deficiencies to putative point mutations in the screen with the translocation would support the hypothesis.

The mapping of the Tc1 induced deficiencies allowed for the division of LGV(left) into more zones. With Stewart *et al.*'s (1990) UV irradiation induced deficiencies (*sDf*'s 70,71,72,73,74 and 75), and the resolution of end points of deficiencies through the

analysis of EMS induced putative point muations (Section 4), LGV(left) has been divided into 36 zones (Figure 12).

The majority of the work in this Section has been included in a publication by Clark, Johnsen, McKim and Baillie (1990).

Section 3:

A Model of Pairing for Recombination Based on Recombination in Deficiency Heterozygotes

INTRODUCTION

It was noted by Brenner (1974) that genes with visible phenotypes show distributional clustering as measured by recombination (except possibly for the X-chromosome). A similar distribution has been observed for the essential genes [see Genomic Maps by Edgley and Riddle (1990); and Section 4 for LGV(left)], where there are regions of gene clustering and regions with few genes. This recombinational distribution may not reflect the physical spacing of genes for two reasons: 1) Results from identifying the number of coding elements in regions of known length both physically and recombinationally suggest that gene spacing along the chromosome and recombinational distances do not have a strong correlation (Greenwald *et al.* 1987; Prasad and Baillie 1989; Starr *et al.* 1989). 2) Recombination distances measured in *C. elegans* males and hermaphrodites show sex specific regional differences (M. Zetka pers. comm.). Recombination distances can be modified. X-rays have been shown to expand recombination in the cluster near the *unc-13* region (Kim and Rose 1987). Temperature and parental age (Rose and Baillie 1979a) as well as the mutation *rec-1* (Rose and Baillie 1979b; Rattray and Rose 1988) all affect recombination.

The decoupling of physical and recombinational gene distances occurs in other organisms (Szauter 1984; for a Drosophila review see Baker *et al.* 1976; for Yeast see Lambie and Roeder 1986).

Studies of translocations and duplications have shown that one half of each *C. elegans* chromosome is essential for disjunction and recombination Rosenbluth and Baillie (1981). Furthermore, the behavior of translocations can be explained by proposing that a single homolog recognition site exists McKim *et al.* 1988b. In addition to chromosomal sites, Hodgkin *et al.* (1979) showed that some mutations that affect disjunction also affect recombination.

In heterozygous translocation strains, recombination is eliminated in the translocated regions. An example of this is *eT1(III;V)* (Rosenbluth and Baillie 1981), a reciprocal translocation of the right half of LGIII and the left half of LGV. In heterozygotes, recombination is suppressed on LGIII(right) and LGV(left), whereas in an homozygous *eT1*, recombination is not suppressed (R.E. Rosenbluth pers. comm.). The recombination suppression may be caused by a failure of the homologous chromosome arms to pair (Rosenbluth and Baillie 1981) or by a failure of homologues to come together Rose and McKim (1989).

McKim and Rose (1990) analyzed 76 duplications of LGI. They noted that the stability of free duplications is not necessarily correlated with size, but may be influenced by specific sequences along the chromosome. Their results strongly suggest that there are sequences spaced along the chromosome that are needed for mitotic pairing. Sequences similar to these, but used meiotically, have also been postulated to exist in Drosophila and have been called pairing points (Hawley 1980). Rosenbluth *et al.* (1985; 1988) noted the phenomenon of recombination suppression between several mutations (including some deficiencies) to the left of *unc-46* and *unc-46* on LGV(left). This strongly suggested that the heterozygous removal of pairing points affected recombination.

This section discusses the work of a collaborative study published by R.E. Rosenbluth, R.C. Johnsen and D.L. Baillie (1990) where the effects of deficiency heterozygosity on recombination along LGV(left) were analyzed.

MATERIALS AND METHODS

(i) *Map positions*

Figure 8 shows the relationship of LGV(left) relative to the whole LGV chromosome. The positions of genes and deficiencies on LGV(left) are shown in Figure 9 (Rosenbluth *et al.* 1988; Edgley and Riddle 1990; Section 1; Section 2; and Section 4). The position of *sDf74* was mapped by Stewart *et al.*(1990). The deficiencies had been isolated as lethal mutations on *unc-46(e177)V* marked chromosomes and were maintained balanced over *eT1(III;V)*.

(ii) *Recombination measurements*

Since recombination rates have been shown to be temperature dependent (Rose and Baillie 1979a), all F_1 heterozygotes were raised at 20C.

Recombination measurements for Table 3 and Table 4: Appropriate P_0 hermaphrodites were crossed with wild-type N2 males; individual phenotypically wild-type F_1 hermaphrodites were picked and all the mature F_2 progeny of a few F_1s with the desired genotypes were scored.

Recombination measurements for Table 5: To avoid picking a large number of F_1s that did not have the desired genotype, deficiency bearing male strains were constructed with the genotypes:

+/unc-60(s1331) eT1(III); sDf unc-46/eT1(V)

or *+/dpy-11(s287) eT1(III); sDf unc-46/eT1(V)*

or *+/eT1(III); sDf unc-46/unc-42 eT1(V)*.

These were crossed to appropriate P_0 hermaphrodites; wild-type F_1 hermaphrodites were picked and the F_2 progeny from the correct F_1s were scored.

Figure 8

Relationship of LGV(left) to the whole LGV chromosome.

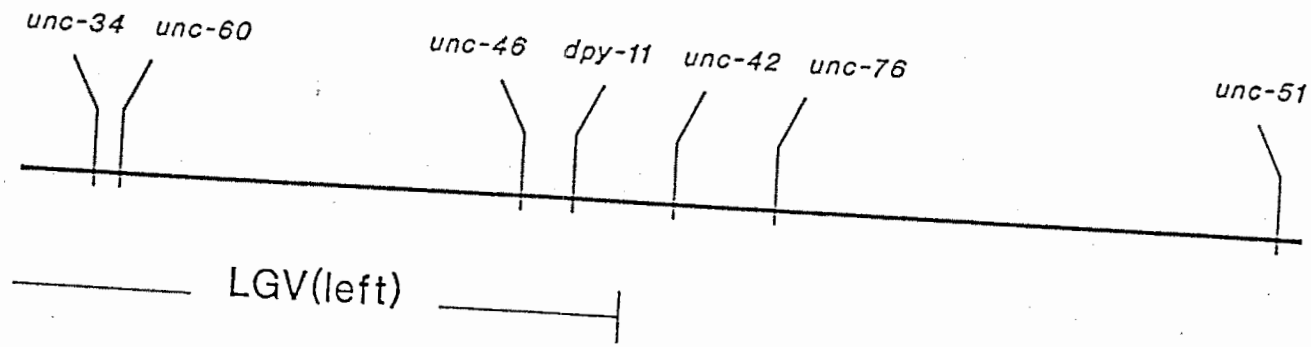
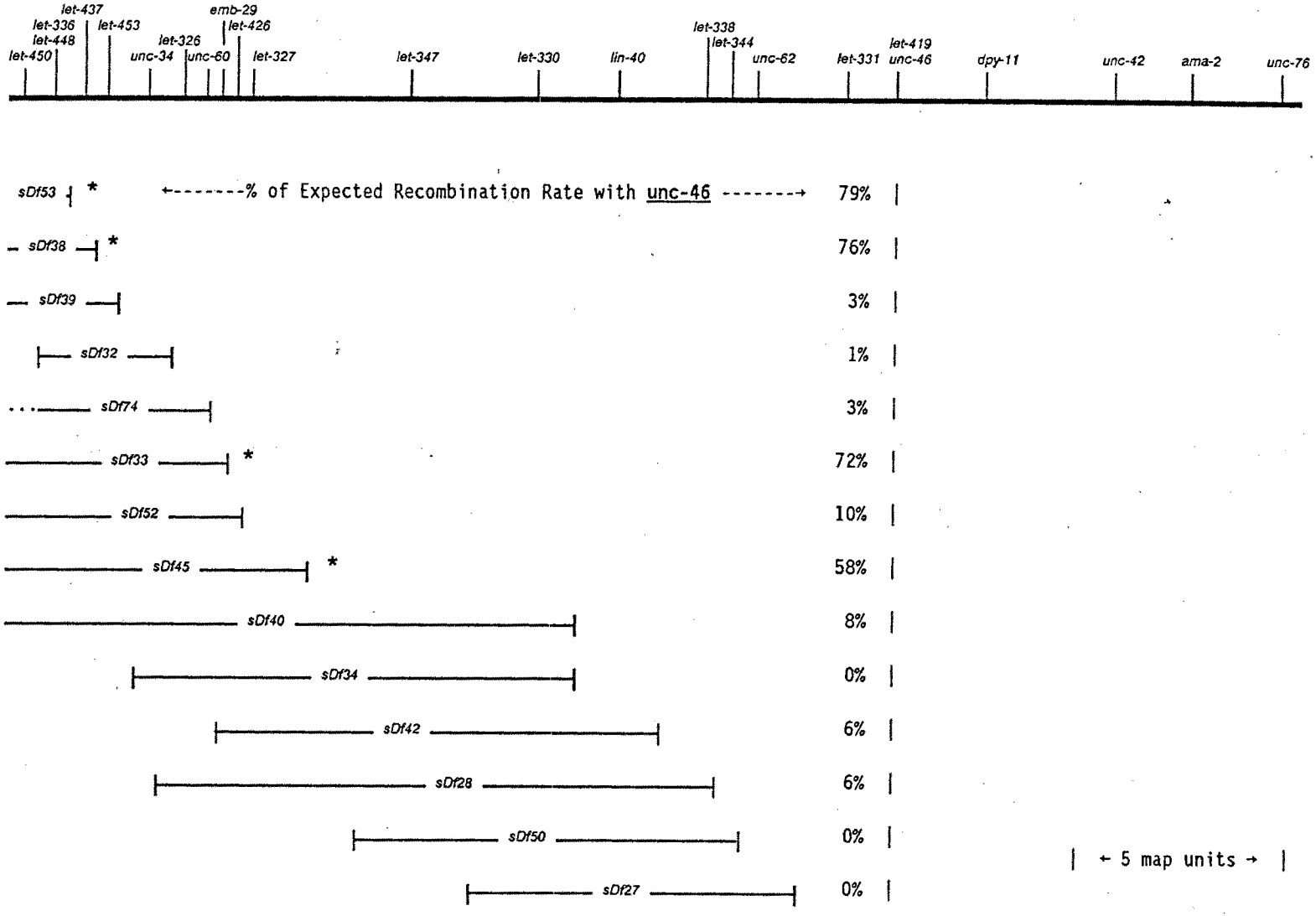

unc-34
unc-60
unc-46
dpy-11
unc-42
unc-76
unc-51
LGV(left)

Figure 9

Percent of expected recombination rate between *Df* and *unc-46*. Measured in *dpy-18/+;Df unc-46/++* hermaphodites. Minor inhibitors are marked with a *. The map distances between *let-448* and *unc-60* are exaggerated for clarity. The map distance between *let-450* and *let-448* is not known.


let-437
let-336
let-448
let-450
let-453
let-326
unc-34
emb-29
let-426
unc-60
let-327
let-347
let-330
lin-40
let-338
let-344
unc-62
let-331
let-419
unc-46
dpy-11
unc-42
ama-2
unc-76
sDf53
*
←--------% of Expected Recombination Rate with unc-46 -------→
79%
sDf38
*
76%
sDf39
3%
sDf32
1%
sDf74
3%
sDf33
*
72%
sDf52
10%
sDf45
*
58%
sDf40
8%
sDf34
0%
sDf42
6%
sDf28
6%
sDf50
0%
sDf27
0%
| ← 5 map units → |

RESULTS

LGV(left) deficiencies fell into two groups: "major" and "minor" inhibitors of recombination: The breakpoints of the 14 deficiencies shown in Figure 9 had previously been localized by complementation mapping (see Materials and Methods for references). All the deficiencies are to the left of *unc-46* and each deletes at least five genes. *sDf53* deletes two more genes than shown (Section 4).

We obtained recombination rates between *unc-46* and the right-hand breakpoint of each deficiency, and compared these rates with those predicted. Table 3, column 4, gives the recombination rates in m.u. and column 6 expresses these in percent of the predicted rates. The latter were based on positions of markers near the respective breakpoints (see Table 4). Figure 9 summarizes the results (expressed as % of the expected rate). Ten of the 14 deficiencies severely inhibited recombination. Each recombined with *unc-46* at less then 11% of the expected rate despite the fact that some deficiencies were at least 17 "m.u." from *unc-46*. These deficiencies will be refered to as "major inhibitors", in contrast to the four "minor inhibitors", *sDf53*, *sDf38*, *sDf33* and *sDf45*. The apparent failure of *sDf34*, *sDf50* and *sDf27* to be able to recombine with *unc-46* (Table 3) has subsequently been shown not to be true. All three deficiencies do recombine with *unc-46* but at very low frequencies relative to the expected ones.

To determine how far inhibition extended along LGV, effects of heterozygous deficiencies were measured across different intervals. The results are summarized in Table 5 and illustrated for the ten major inhibitors in Figure 10. For most deficiencies the distances could be measured only in intervals to the right of the deficiency. In two cases it was possible to measure recombination to the left (from *unc-34* to *unc-60* for *sDf42*; and from *unc-60* to *let-347* for *sDf27*).

Two features of the results were significant. 1) The similarity with which recombination in specific intervals to the right was affected by the major inhibitors, despite the fact that their right-hand breakpoints varied considerably: Some were as much as 18 m.u. from *dpy-11* (*sDfs 39* and *32*) while others were only 5-6 m.u. away (*sDfs 28, 50* and *27*). Inhibition by all ten deficiencies reduced recombination to less than 16% of control in the *unc-62 - dpy-11* interval and the inhibition effectively eliminated in the *ama-2 - unc-76* interval. 2) The recombination suppression was directional. Recombination could be measured on both sides of two deficiencies (*sDf27* and *sDf42*). Both *sDf27* and *sDf42* inhibited recombination to their right but showed no corresponding inhibition of recombination to their left. This biased inhibition could also be demonstrated for *sDf50* by comparing recombination between the *unc-60 - sDf50* and the *unc-62 - dpy-11* intervals. Hermaphrodites with the genotype *unc-60(m35) + unc-62(s472) + dpy-11/+ sDf50 + unc-46 +* produced 61 Unc-60s and 8 Dpy-11 among 2336 adult progeny (*s472* is a recessive lethal allele of *unc-62*). The Dpy-11s were due to recombination between *unc-62* and *dpy-11*, giving an apparent map distance of 0.3 m.u. for this interval instead of the normal 5.3. Thus on the right, *sDf50* heterozygosity decreased recombination to 6% of the normal frequency. The Unc-60s were due to recombination between *unc-60* and *unc-62*, occurring both to the left and to the right of *sDf50* (Figure 10). They represented an apparent recombination distance of 2.6 m.u. Only a negligible part of this recombination could have been to the right [6% of the 0.5 m.u. *let-344 - unc-62* distance (Figure 10)]. Therefore, the 2.6 m.u. represented recombination between *unc-60* and the left breakpoint of *sDf50*. Since the actual distance could be no more than 4.1 m.u. (the normal *unc-60 - let-347* distance), recombination to the immediate left of *sDf50* was at least 63% of the normal frequency, which is in sharp contrast to 6% on the right.

Table 3
(Recombination between unc-46 and deficiencies[a])

Deficiency	Unc-46 recombinants	F_2 adults	recombination in map units		
			Actual[b]	Expected	% of expected
*sDf53**	134	2149	13.4(11.0-15.9)	>17	<79
*sDf38**	132	2172	13.0(10.7-15.6)	>17	<76
sDf39	3	2316	0.5(0.1- 0.7)	>17	< 3
sDf32	1	2300	0.1(0.0- 0.5)	~16	~ 1
sDf74	2	1133	0.4(0.1- 1.2)	~16	~ 3[c]
*sDf33**	120	2344	10.8(8.9-13.0)	~15	~72
sDf52	10	1305	1.5(0.8- 2.7)	~15	~10
*sDf45**	95	3067	6.4(5.1- 8.0)	>11	<58
sDf40	3	1294	0.5(0.1- 1.3)	> 6	< 8
sDf34	0	2263	0.0(0.0- 0.3)	> 8	< 0
sDf42	2	2174	0.2(0.0- 0.6)	> 3	< 6
sDf28	1	3074	0.2(0.0- 0.3)	~ 3	~ 6
sDf50	0	1660	0.0(0.0- 0.4)	> 2.5	0
sDf27	0	2619	0.0(0.0- 0.3)	> 1.2	0

* "Minor" inhibitors (see text).

a Data taken in part from Rosenbluth *et al.* (1988); Clark *et al.* (1990).

b 1 map unit = 100p, where $p=1-[1-4(U46)]^{1/2}$, where U46 is the frequency of Unc-46 recombinants. 95% confidence limits, in brackets, are based on limits of the recombinants which are taken from Table 1 of Crow and Grardner (1959).

c Data from Stewart *et al.* 1990.

Table 4
[Two-factor recombination data for genes on LGV(left)]

Recombination interval	P_0 hermaphrodites	F_1 progeny: Recombinants[a]	F_1 progeny: Total	Equation for p[b]	Distance in map units[c]
let-448 to *unc-46*	*dpy-18/+;let-448 unc-46/++*	105 U46	1415	[1]	16.1(12.8-19.8)
unc-34 to *unc-60*	*unc-34 unc-60(m35)/++*	22 U34	3896	[1]	1.1(0.7- 1.7)
let-326 to *unc-46*	*dpy-18/+;let-326(s238) unc-46/++*	126 U46	1592	[1]	17.3(14.1-21.0)[d]
let-326 to *unc-46*	*dpy-18/+;let-326(s1404) unc-46/++*	265 U46	3660	[1]	15.7(13.7-17.9)
unc-60 to *emb-29*	*unc-60(e677) emb-29 dpy-11/+++*	7 U60	2488	[2]	0.4(0.2- 0.8)[e]
unc-60 to *let-347*	*unc-60(m35) let-347 dpy-11/+++*	111 U60	4192	[2]	4.1(3.3- 4.8)
unc-60 to *dpy-11*	*unc-60(m35) dpy-11/++*	647 U60&D11	3979	[3]	17.8(16.7-19.0)
let-327 to *unc-46*	*dpy-18/+;let-327 unc-46/++*	184 U46	2683	[1]	14.8(12.5-17.3)[d]
let-347 to *dpy-11*	*unc-60(m35) let-347 dpy-11/+++*	339 D11	4192	[2]	13.0(11.7-14.3)
let-330 to *unc-46*	*dpy-18/+;let-330 unc-46/++*	37 U46	943	[1]	8.3(5.7-11.3)[d]
lin-40 to *dpy-11*	*lin-40(e2173) dpy-11/++*	97 D11	2391[f]	[1]	8.5(6.8-10.3)[d]
let-338 to *unc-46*	*dpy-18/+;let-338 unc-46/++*	74 U46	4351	[1]	3.5(2.7- 4.3)[d]
let-344 to *unc-46*	*dpy-18/+;let-344unc-46/++*	18 U46	1276	[1]	2.9(1.8- 4.4)[d]
unc-62 to *unc-46*	*dpy-18/+;unc-62 unc-46/++*	26 U46	1925	[1]	2.7(1.8- 4.0)[d]
unc-62 to *dpy-11*	*unc-62 dpy-11/++*	102 D11	3046	[2]	5.2(4.2- 6.2)
let-331 to *unc-46*	*+ let-331 unc-46 +/unc-60 ++ dpy-11*	7 U46	894	[2]	1.2(0.6- 2.3)[d]
unc-46 to *dpy-11*	*unc-46 dpy-11/++*	71 U46&D11	3337	[3]	2.1(1.7- 2.7)

a Abbreviations for phenotypes: U46 = Unc-46; U34 = Unc-34; U60 = Unc-60; D11 = Dpy-11.

b Equations for the recombination frequency, p, where R = frequency of recombinants scored:

[1] $p = 1(1-4R)^{1/2}$

[2] $p = 1(1-3R)^{1/2}$

[3] $p = 1(1-2R)^{1/2}$

c 1 map unit = 100p. The 95% confidence limits (in brackets) are based on the limits of the recombinants. These limits are either taken from Table 1 of Crow and Gardner (1959), or, $= 2[Nq(1-q)]^{1/2}$ where N = total F_1s and q = frequency of recombinants (for >300 recombinants).

d Data taken from Rosenbluth *et al.* (1988).

e Data taken from McKim *et al.* (1988a).

f Includes larval F_1s

Table 5
(Effects of deficiency heterozygotes on crossing over in LGV intervals)

Deficiency	Percent recombination of control value in different intervals[a]						
	unc-34-unc-60	*unc-60-let-347*	*unc-62-dpy-11*	*unc-46-dpy-11*	*dpy-11-unc-42*	*unc-42-ama-2*	*ama-2-unc-76*
Control	100.0(63.6-154.5)	100.0(81.4-123.3)	100.0(84.2-116.1)	100.0(71.2-137.2)	100.0(82.1-117.8)	100.0(75.1-124.9)	100.0(72.0-134.7)
*sDf53**	ND	72.1(55.8-93.0)	86.8(67.9-103.8)	ND	ND	ND	ND
*sDf38**	ND	62.8(46.5-76.7)	75.5(60.4- 92.5)	ND	ND	ND	ND
sDf39	ND	2.3(0.0-11.6)	7.5(3.8- 13.2)	ND	23.2(13.7-32.7)	75.2(50.3-100.0)	128.6(95.2-176.2)
sDf32	#	0.0(0.0- 4.7)	5.7(3.8- 11.3)	ND	0.0(0.0- 2.0)	126.6(84.2-175.7)	128.6(85.7-181.0)
sDf74	#	#	15.1(7.5- 22.6)	ND	52.6(40.4-66.6)	ND	133.3(95.2-176.2)
*sDf33**	#	#	71.7(48.8- 95.7)	ND	ND	ND	ND
sDf52	#	#	15.1(7.5- 22.6)	ND	34.4(23.8-47.4)	ND	50.5(30.5-80.9)
*sDf45**	#	#	34.0(20.8- 49.1)	ND	ND	ND	ND
sDf40	#	#	3.8(1.9- 11.3)	ND	16.2(10.3-23.5)	ND	100.0(61.9-142.9)
sDf34	#	#	3.8(1.9- 11.3)	ND	8.3(4.6-20.8)	57.1(37.3-77.4)	81.0(57.1-104.8)
sDf42	118.2(72.7-172.7)	#	13.2(5.7- 20.8)	ND	25.2(17.2-35.1)	113.0(83.1-149.7)	66.7(47.6- 85.7)
sDf28	#	#	3.8(1.9- 9.4)	ND	13.9(7.6-23.5)	65.5(42.4-97.2)	104.8(71.4-147.6)
sDf50	ND	#	5.7(1.9- 11.3)	ND	18.0(10.7-25.7)	100.6(70.1-131.2)	ND
sDf27	ND	104.7(86.0-132.6)	#	21.9(9.8-47.0)	43.0(33.1- 53.0)	ND	95.2(71.4-119.0)

* "Minor" inhibitor.
ND Not determined.
\# Deficiency extends into the interval.
[a] 95% confidence limits are in brackets.

DISCUSSION

Recombination on LGV(left) is extensively suppressed by deficiencies. A set of 14 deficiencies were analyzed with respect to their effects on recombination. The fourteen deficiencies include six internal deficiencies and eight deficiencies that delete the left-most known gene on LGV (*let-450*). All the deficiencies were induced on chromosomes marked with *unc-46* which lies to the right of all 14 deficiencies. There are two classes of deficiencies that inhibit recombination: 1) Major inhibitors that suppress recombination almost completely (>90% between the right breakpoint of the deficiences and *unc-46*). 2) Minor inhibitors that suppress recombination to a minor degree (20-50% between the deficiencies and *unc-46*). All the internal deficiencies are major inhibitors but some of the deficiencies that delete the left-most known gene (*let-450*) are also major inhibitors. All minor inhibitors delete *let-450*.

The inhibition cannot be attributed to the deletion of a single site because there is no common point deleted by all the major inhibitors. The extent of inhibition is independent of the right-hand deficiency breakpoints. The amount of inhibition of recombination is similar for all the major inhibitors in all the intervals tested. The inhibition of recombination decreases with distance and disappears in the *ama-2 - unc-76* interval.

There is a polarity to the inhibition of recombination. Where inhibition of recombination could be tested to the left of the deficiencies (*sDf27*, *sDf42* and *sDf50*) there was no evidence for any suppression.

To explain the differences between the minor and major inhibitors, a site is postulated to exist near the left terminus of the chromosome that is deleted by the minor inhibitors but not by the major inhibitors. The six major inhibitors that are known to be internal deficiencies do not delete this site. Even though the other four major inhibitors

delete *let-450*, it is postulated that they are internal deficiencies. Since it is not known have far to the left of *let-450* LGV extends it is plausible that the four deficiencies are internal.

To explain the above results, a model for the manner in which pairing for recombination takes place has been erected. The model proposes that there are two types of sites needed for pairing for recombination; "initiation sites" and "alignment sites". Intimate pairing for recombination starts at and spreads from each initiation site, if homologous sites are within a minimum distance of each other. Initiation sites are analogous to the pairing points in *D. melanogaster* mapped by Hawley (1980). The special site that is postulated to exist near the left terminus of LGV is an initiation site. Another site must exist in LGV(right) (Figure 11A) to account for recombination occurring in that region in *eT1(III;V)* translocation heterozygotes. The pairing process spreads from each initiation site by "buttoning-up" the homologs at sequential alignment sites that occur repeatedly between initiation sites and have a common sequence. In the presence of a heterozygous minor inhibitor, there are no homologous left-end initiation sites. Pairing initiates only in LGV(right) (Figure 11B), proceeds toward the left, aligns the homologs in a correct manner (*i.e.*, homologous regions remain in register) and no major inhibition of recombination occurs. On the other hand, in the presence of a heterozygous major inhibitor (which does not delete the initiation site), pairing proceeds from both ends of LGV (Figure 11C). To the left of the deficiency, homologous regions remain in register (*i.e.* region "a") and recombine normally. But to its right, the sequential buttoning of alignment sites cause a misalignment: heterologous regions become aligned, thus inhibiting recombination. This misalignment continues until a region is reached whose alignment is controlled by pairing that initiated at the right end. To explain the disappearance of inhibition in the *ama-2* region, it is proposed that the speed at which pairing occurs from each end is such that alignment of the *ama-2* region is mainly controlled from the right and

therefore remains normal, while the alignment of the *unc-62* region (on the left) is still controlled by pairing that initiated at the left end and is not normal. Based on the current data, there would be a minimum of three alignment sites in LGV(left): One in the *sDf32* region, one in the *let-330* region and one between *sDf27* and *dpy-11*. Finding small deficiencies within these three regions would help localize the alignment sites more precisely. Alternatively, new LGV(left) deficiencies lying outside these regions would either identify more alignment sites or indicate that the deficiencies lie between two such sites.

Translocation and duplication experiments implicate regions for each of five chromosomes as homologous recognition sites necessary for segregation and recombination (referenced in Introduction). From the results of Rosenbluth and Baillie 1981; Herman *et al.* 1982; and Ferguson and Horvitz 1985, A. Rose (Pers. comm.) argued that the homologous recognition site on LGV is near the right end. This is not inconsistent with the above model where there is a proposed initiation site for pairing at the left end of LGV. What the model requires is that the initiation sites for pairing for recombination act secondarily to the homolog recognition site.

If the above model is correct then it is reasonable to use it to speculate on the recombinationally non-uniform spacing of genes along the chromosomes. The proposed alignment sites can be envisioned to be non-uniformly spaced along the chromosome. This would give them one of two characteristics necessary to account for the uneven recombinational spacing. The other necessary characteristic is that the alignment sites can be recognized by the recombination machinery and therefore used to (roughly) localize recombination. Using short deficiencies it should be possible to identify the location of alignment sites. If the alignment sites determine recombination then the spacing of the alignment sites should be constant in terms of map units. Stewart *et al.* (1990) has identified an internal deficiency (*sDf75*) that is not a major inhibitor of recombination (for

location see Figure 12). A crude estimate (from the data in Table 4) of the length of chromosome deleted by *sDf75* is 2 m.u.. Using this estimate, I propose that alignment sites are at least 2 map units apart.

Figure 10

Percent of control recombination in different LGV intervals, in *Df* heterozygotes. The effects of only the major inhibitors are shown.

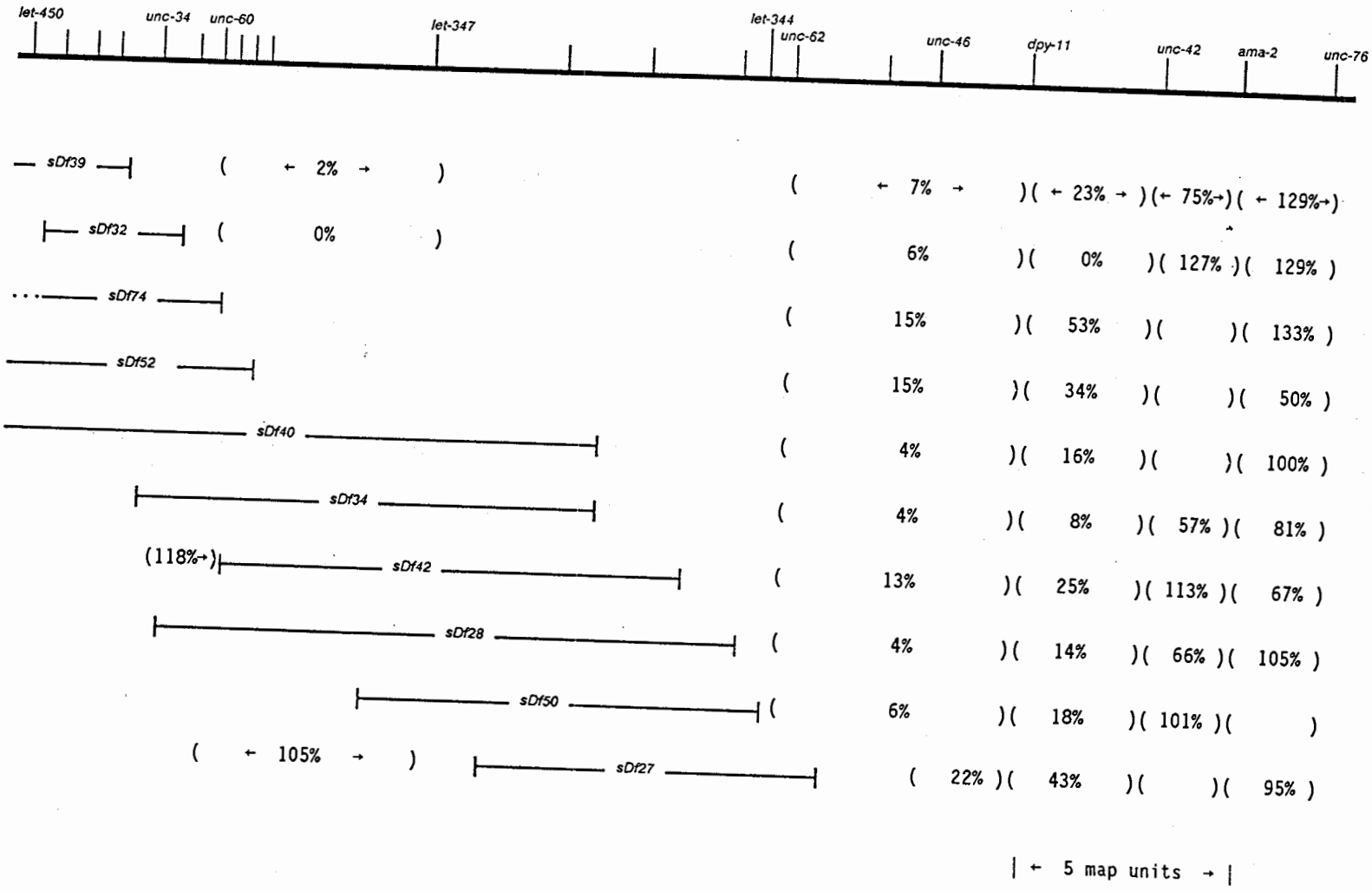
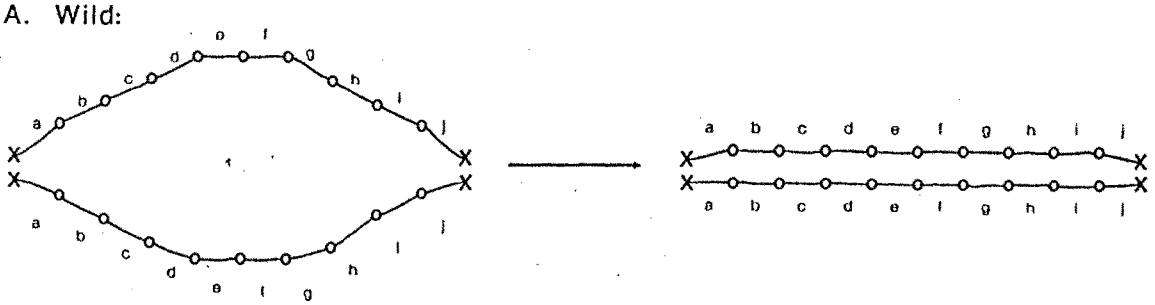

let-450
unc-34
unc-60
let-347
let-344
unc-62
unc-46
dpy-11
unc-42
ama-2
unc-76
62 B
sDf39
(← 2% →)
sDf32
(0%)
sDf74
sDf52
sDf40
sDf34
(118%→)
sDf42
sDf28
sDf50
(← 105% →)
sDf27
(← 7% →)(← 23% →)(← 75%→)(← 129%→)
(6%)(0%)(127%)(129%)
(15%)(53%)( )(133%)
(15%)(34%)( )(50%)
(4%)(16%)( )(100%)
(4%)(8%)(57%)(81%)
(13%)(25%)(113%)(67%)
(4%)(14%)(66%)(105%)
(6%)(18%)(101%)( )
(22%)(43%)( )(95%)
| ← 5 map units → |

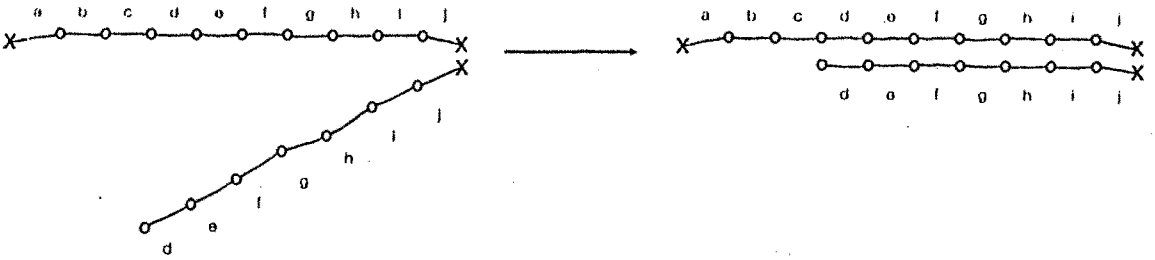
Figure 11

Proposed pairing for recombination along LGV. a, b, c,...j are chromosomal regions.

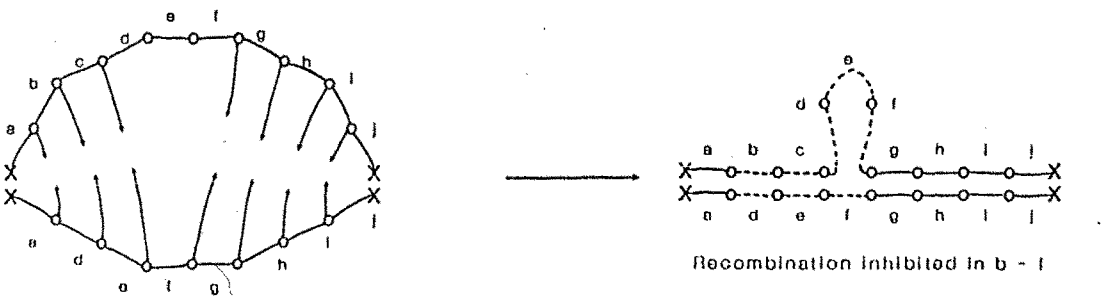
A. Wild:



B. Heterozygote for a - c deficiency:



C. Heterozygote for b - c deficiency:



Recombination inhibited in b - i

X = Site at which pairing for recombination initiates.

o = Alignment site.

---- = Regions in which recombination is inhibited.

Section 4:

Towards saturation of LGV(left) for essential genes

INTRODUCTION

In this section I have attempted to identify the majority of essential genes on the 23 m.u. of LGV(left) that is balanced by the reciprocal translocation *eT1* (Rosenbluth 1981). LGV(left) contains approximately 7% of the recombinational distance in *C. elegans'* genome. LGV(left) has been subdivided into zones by sets of overlapping deficiencies (Johnsen and Baillie 1988; Rosenbluth *et al.* 1988; Clark *et al.* 1990; Stewart *et al.* 1990). At the start of this Section, 17 putative non-essential genes (under laboratory conditions) and 63 essential genes had been identified on LGV(left) [Johnsen and Baillie 1988 (Section 1); Rosenbluth *et al.* 1988; Clark *et al.* 1990 (Section 2); Edgley and Riddle 1990; Stewart *et al.* 1990].

To identify the majorities (approximately 70%) of the essential genes on LGV(left), one must be able to induce mutations within genes. EMS is a monofunctional alkylating agent specific for G-C to A-T transitions and therefore useful for inducing point mutations (Discussed in Fincham 1983). To identify the essential genes one must know how many 0.012M EMS (dose recommended by Rosenbluth *et al.* 1983) treated chromosomes to screen. Rosenbluth *et al.* (1988) had screened approximately 6000 EMS treated *eT1* balanced chromosomes. They recovered recessive lethal alleles of 41 genes, and for 10 of those genes they recovered more than one allele. Using a truncated Poisson calculation (as described by Meneely and Herman 1979) yielded a mean number of mutations per essential gene *(m)* of 0.54 and an estimate that there are a minimum of 98 essential genes on LGV(left). The dosages of EMS used by Rosenbluth *et al.* (1988) ranged from 0.004-0.058M. They recovered approximately 480 recessive lethal mutations balanced by *eT1*. The forward mutation rate of 0.012M EMS was reported to be 6.6% over the *eT1* balanced region (Rosenbluth *et al.* 1983). My intent was to increase the number of identified genes to at least 70% of all the essential genes on LGV(left). This required an *m*

of approximately 1.2 and therefore the generation and analysis of 720 *eT1* balanced recessive lethal mutations. To do this, approximately 11,000 chromosomes treated with 0.012M EMS had to be screened.

I have screened 10,900 *eT1* balanced chromosomes treated with 0.012M EMS and isolated 751 recessive lethal (including sterile adult and some *mel*) mutations in or close to the balanced region. I compared the number of lethals that mapped to each of the two *eT1* balanced chromosomes. I used complementation tests to map the lethals on LGV(left) against a set of deficiencies (Rosenbluth *et al.* 1988; Section 1; Section 2) that divide LGV(left) into zones, and once the lethals were mapped to zones I used complementation tests to map them against appropriate genes to discern if the lethals identified new genes or were alleles of known genes. I noted the approximate blocking stages for each lethal at 20C and tested to see if any of the mutations were rescuable at 15C or 25C. I also tested the sterile adults and *mels* for male rescuability.

MATERIALS AND METHODS

EMS mutagenesis:

Mutagenesis of BC2200 [*dpy-18/eT1(III);unc-46/eT1(V)*] was done according to Brenner (1974) except that 0.012M EMS was used as recommended by Rosenbluth *et al.* (1983). Mutagenesis was done at room temperature (approximately 21-23C).

Screening for mutations:

After mutagenesis, the egg bearing young adult P_0s were allowed to recover for two hours and then were individually plated and brooded for 21 hours ("A" brood) at room temperature, then 45 hours at 15C ("B" Brood), the P_0s were then discarded. F_1s were picked individually (for every P_0 used, all F_1s were set up). The F_1's progeny were screened for the absence of healthy egg-bearing Dpy Uncs (lethal, sterile and some *mel* mutations were picked up). The *dpy-18/eT1(III);unc-46/eT1(V)* hermaphrodite's progeny are in a 1 Dpy Unc : 4 wild-type : 1 eT1(Unc-36 phenotype) ratio, all aneuploids die as eggs or L1 larvae (L. Turner pers. comm.). The absence of fertile Dpy Unc F_2s indicated the presence of a lethal mutation in the *eT1* balanced regions of LGIII or LGV (or outside but close to the balanced region).

Developmental blocking stage

All developmental blocking stages were noted at 20C. Egg laying was noted for Dpy-18 Unc-46 after they reached 0.7mm. They reach a mature length of 0.8-0.9mm. A growth curve noting moulting stages has not been done. Prior to analysis, the worms were set at 20C for at least one generation to eliminate any possible temperature dependent maternal effects. Worms containing relatively early blocking recessive lethals were handled as follows: *dpy-18/eT1(III);let unc-46/eT1(V)* hermaphrodites were mated to BC1958 *dpy-18/eT1(III);unc-46/eT1(V)* males. One to three Dpy Unc F_1s were brooded for 6 hours and the number of eggs (F_2s) laid were counted. The next day the F_2s were

scored for the presence of unhatched eggs (putative egg lethals). On the third day the F_2s were counted and the maturing phenotypically wild-type F_2s removed. On the fourth and seventh days the homozygous lethal bearing Dpy Uncs were observed and measured with a micrometer. Worms containing late larval to *mel* mutations were handled as follows: approximately 20 *dpy-18/eT1(III);let unc-46/eT1(V)* hermaphrodites were brooded for 6 hours. The Dpy Unc progeny were transferred to new plates for observation and measurement on the fourth and seventh days.

The homozygous lethal containing Dpy Unc worms were classified according to their lengths. Worms shorter than 0.2mm were considered to have not developed beyond hatching, worms between 0.2 and 0.3mm were called early blockers, 0.3-0.5mm mid blockers, 0.5-0.6 late blockers and greater than 0.6mm with no internal fertilized eggs were called sterile adults. Worms with internal fertilized eggs or developmentally blocked progeny were considered to be homozygous for *mel* mutations. Worms that were fertile but slow developing were noted and called "slow".

Temperature rescue

Worms were set at either 15C or 25C for at least one generation before being tested. *dpy-18/eT1(III);let unc-46/eT1(V)* hermaphrodites were brooded at either 15C or 25C and their progeny scored for the presence of Dpy Uncs. If Dpy Uncs were present, then they were brooded individually and scored for fertility.

Male rescue

Sterile adult and *mel* mutations were tested for male rescue by mating 3 Dpy Uncs containing the homozygous mutation, with 3 N2 males. The presence of wild-type male and fertile wild-type hermaphrodite progeny indicated male rescue had occurred. The absence of fertile wild-type progeny indicated either no male rescue or, in the cases of hermaphrodites with vulva abnormalities, most likely failure to mate.

RESULTS

A total of 10,900 chromosomes treated with 0.012 M EMS were screened for the presence of lethal mutations on LGIII(right) or LGV(left). 751 mutations were isolated: 242 strains contained mutations strictly on LGV, 451 strictly on LGIII and 29 strains carried mutations on both LGV and LGIII. All strains but the ones carrying mutations strictly on LGV were discarded.

When screening for lethal mutations in the *eT1* balanced regions, mutations were expected to be recovered that were close to but not in the *eT1* balanced region. While analyzing the 242 strains containing mutations on LGV(left), 50 (27%) of the mutations were lost and therefore presumed to have been outside the balanced region. Two other strains proved to have two lethal mutations within the *eT1* balanced region on LGV(left). Therefore, a total of 194 mutations in essential genes in the *eT1* balanced region of LGV were recovered. This gave a 2% forward mutation rate for 0.012 M EMS in the *eT1* balanced region of LGV.

The *eT1* breakpoints on both LGIII and LGV are in gene clusters and the *eT1* breakpoint on LGIII is in or very close to *unc-36* [Rosenbluth and Baillie 1981; See Edgley and Riddle (1990) for the genetic map of LGIII]. Assuming that the same number (50) of mutations would have been outside the *eT1* balanced region on LGIII as were lost from LGV, then the forward mutation rate for 0.012M EMS is 5.8% over the region balanced by *eT1*. This is in agreement with Rosenbluth *et al.* (1983), where the screening of 1662 chromosomes yielded a forward mutation rate of 6.6%.

Of the 10,900 chromosomes screened for the presence of lethal mutations, 4750 (43.6%) came from the A brood and 6150 (56.4%) came from the B brood. Of the 194 mutations recovered on LGV(left), 84 (43.3%) came from the A brood and 110 (56.7%)

came from the B brood. Therefore there is no significant difference in the mutation rates of the early and late broods.

There were no "jackpots" from this screen, no gene had more than one allele originating from any given P_0, *i.e.* there was no evidence for premeiotic events. 166 of the 194 lethal mutations were mapped, the other 28 were difficult to work with (*i.e.* male strains could not be established or maintained) and therefore not mapped. The analysis of the 166 (my EMS set) recessive lethal mutations yielded one deficiency (*sDf56* which deletes zones 1-4B inclusive, see Figure 12), and alleles of 74 genes including 38 new essential genes and the first lethal allele of *unc-60*. A total of 101 genes with recessive lethal alleles have now been identified on LGV(left). Rosenbluth *et al.* (1988) screened for recessive lethals using EMS mutagenesis. They analyzed 54 mutations recovered from screens for recessive lethal mutations over the entire *eT1* balanced region of LGV(left). Combining their 54 mutations with my EMS induced mutations gave 220 mutations (Total EMS set) in 89 of the 101 genes. Mutations defining the other 12 genes came from screens for selected regions of LGV(left) or from screens using mutagens that can cause rearrangements (gamma irradation and formaldehyde) or from a mutagen that has preferred sites of action (mobilization of Tc1). Figure 12 is a genetic map of LGV(left). The genes are above the line representing the chromosome; and the rearrangements [deficiencies and one duplication (*sDp30*)] that divide the region into zones are below. The zones are listed at the bottom.

Figure 13 shows the distribution of the number of genes with given numbers of alleles from the total EMS set. The Poisson distribution in Figure 13 is based on an average hit frequency per gene of 1.25 (see Discussion). *let-330* is the largest target (16 alleles) for EMS induced mutagenesis. *rol-3* and *let-332* are also large targets with 8 alleles each. Table 6 lists the number of alleles for all the essential genes mapped on LGV(left).

In Appendix A, the alleles of the essential genes mapped to LGV(left) are listed according to the screen from which they were isolated. Allele numbers between s1400 and s1640 as well as s2167 and s2168 are from my EMS set. The 38 genes first identified in this section are indicated with an asterisk (note that *ges-1*, *mec-1*, *kar-1* and *unc-83* were not complementation tested against lethals). Two asterisks indicate genes that I have identified in other screens. I have identified a total of 50 genes on LGV(left). The two alleles of *let-466* (zone 8A2) were not from my EMS set, but were analyzed with them. They were isolated from an "eT1" screen discussed by Rosenbluth *et al.* (1988), as such they belong in the total EMS set. The analysis of my EMS set has helped to distinguish the breakpoints of some of the rearrangments that subdivide LGV(left) into zones. LGV(left) is now divided into 36 zones. The 101 essential genes and 17 putative nonessential genes total 118 genes mapped to zones on LGV(left). The average number of genes/zone is 3.3; the number of genes/zone range from one to 12 (both zone 10 and 15 contain 12 genes). Therefore the majority of genes on LGV(left) have been right/left positioned with respect to each other by deficiency mapping. The selection for new rearrangments that break in the zones that contain several genes would be useful in right/left positioning the genes in those zones, but ultimately 3-factor mapping of genes within zones may be needed. Because of the large number of deficiency breakpoints mapped on LGV(left) (see Figure 12) some should break in or near genes of interest. Thus the molecular identification of appropriate breakpoints could facilitate gene cloning.

In Appendix B, the developmental blocking stages of the various alleles are listed (note that outcrossing was not exhaustive and therefore variations of phenotypes among alleles of a given gene may be an artefact due to undetected second mutations). I have changed the canonical alleles for some of the lethal genes previously identified in this laboratory. The canonical alleles were only changed to reflect an earlier blocking recessive lethal phenotype or an EMS induced allele of a gene identified only by a gamma irradiation

induced mutation. The canonical alleles are listed in column 2. The blocking stages of the canonical alleles range from egg lethal to slow developers. The number of genes of each category is shown in Figure 14. Because I screened for recessive lethal mutations, the *mel* class may be underrepresented. Previously I screened (see Rosenbluth *et al.* 1988) 2738 *eT1* balanced chromosomes, and retained all F_1s until the Dpy Uncs were tested for fertility, thus selecting for *mel* mutations. Two *mel* mutations were isolated on LGV(left). In my screen of 10,900 chromosomes, 6 *mel* mutations on LGV(left) were isolated serendipitously, suggesting that about three/quarters of the *mel* mutations generated were isolated. The isolation of these *mel* mutations implies that a large fraction of *mel* mutations have phenotypic effects other than inviable progeny. The reason for this is that *mels* should not have been isolated from my screen. These mutations were probably picked up because they showed some other phenotypic effect (*i.e.* slow growing or thin due to few eggs) and subsequently were shown to be Mel.

I attempted to test all (not just one allele/gene) of my EMS set for temperature rescue at 15C and 25C. It should be noted that the "B" brood P_0s were maintained at 15C before their F_1s were brooded. Some temperature rescued mutations could have been rescued and therefore missed. Table 7 shows the rescuable, partially rescuable and not done (N.D.) mutations. None of the other mutations showed any obvious temperature effects. Worms homozygous for four mutations [*let-422(s1578), let-476(s1621), let-470(s1629)* and *let-441(s1414)*] were rescued at 15C. At 20C, Let-441(*1414*) escapes, [an escaper is defined as an occasional individual, who carries a lethal factor in an effective dose, that overcomes the crisis and develops further (Hadorn 1961 cites Hadorn 1945)] and occasionally a worm is fertile; at 15C, six of nine Dpy-18 Unc-46 *let-441(s1414)* worms tested were fertile; at 25C, the worms were not healthy and none were fertile. Although worms homozygous for *let-461(s1486), let-418(s1617)* or *let-442(s1430)* blocked development at a more advanced stage at 15C than at either 20C or 25C, they were not

rescued (they should be tested at lower temperatures). No mutations were rescued at 25C but two [*s1446*, and *let-343(s1465)*] were partially rescued.

Table 8 has the results from testing homozygous mutations (from my EMS set) that produced sterile adults, maternal effect lethals or slow developers for rescue by wild-type sperm. I also included in Table 8 sterile and *mel* canonical alleles from Rosenbluth *et al.* (1988) although they were not tested. Nine mutations were rescued by male sperm. For one [*let-479(s1576)*], unmated hermaphrodites laid "mushy" unfertilized eggs, but laid apparently normal eggs (that developed into normal worms) when fertilized by wild-type males. This implies that *let-479* is involved in the development or functioning of sperm. *let-346*, *let-469*, *let-475* and *let-476* also have male rescued recessive alleles that otherwise cause adult sterility. These genes could also be involved in the development or functioning of sperm.

Three of the identified essential genes (*let-418*, *let-420* and *lin-40*) have recessive alleles that not only cause adult worms to be sterile but also to have protruding vulvae and thus these genes putatively affect the vulva-lineage.

Four of the 10 genes on LGV(left), that were first identified for their obvious recessive non-lethal phenotype (*i.e.* Dpy, Rol, Unc) now have recessive lethal alleles. They are: 1) the first recessive lethal allele (*s1586*) of *unc-60* (predicted by McKim *et al.* 1988a) came from my EMS set; 2) one spontaneous recessive lethal allele of *unc-62* (Rosenbluth *et al.* 1988); 3) six recessive lethal alleles *unc-70* (predicted by Park and Horvitz 1986a), one from Rosenbluth *et al.* (1988) and five (*s1406*, *s1502*, *s1532*, *s1557* and *s1639*) from my EMS set. Worms heterozygous for most of the recessive lethal alleles of *unc-70* appear to be somewhat uncoordinated. Worms homozygous for most of the alleles tend to coil or and have their development arrested at an early larval stage. Occasionally, worms homozygous for *unc-70(s1502)* escape. They mature and give early larval blocked progeny with an occasional escaper. My lethal alleles were not complementation tested against

e524 (the visible allele). They were tested against *s115* which is an early recessive lethal allele; 4) Twelve recessive lethal allele of *rol-3*. I have identified five (*s1408*, *s1409*, *s1473*, *s1494* and *s1519*) from my EMS set, the other seven were identified by Rosenbluth *et al.* (1988). All of my alleles (with the possible exception of *s1473*) are recessive for an early larval lethal phenotype. *s1473* is in a strain with another recessive lethal mutation [*let-463*(*s2168*)]. The blocking stages due to these mutations individually were not determined. One of the *rol-3* alleles, which I identified in an earlier screen (see Rosenbluth *et al.* 1988), is temperature sensitive (ts). At 20C it is a recessive lethal but at 15C it is a fertile roller. My lethal alleles were complementation tested against *s442* and/or *s126* (both recessive lethals) but not against the visible *e754*.

Table 6
[Number of alleles of essential genes on LGV(left)]

ZONE	GENE	TOTAL NUMBER OF ALLELES	EMS ALLELES INDUCED OVER *eT1*
1A	*let-450*	1	0
20A	*let-460*	1	0
11A2	*let-401*	1	0
16	*let-406*	1	0
15	*let-438*	1	0
15	*let-404*	1	0
19	*let-416*	1	0
1B2	*let-448*	1	0
16	*let-449*	1	0
17	*let-414*	2	0
9B	*unc-62*	2	0
17	*let-408*	2	0
6B	*let-347*	1	1
11B4	*let-418*	1	1
15	*let-425*	1	1
9A	*let-430*	1	1
17	*let-436*	1	1
1B3	*let-437*	1	1
20B	*let-441*	1	1
1B1	*let-458*	1	1
14	*let-471*	1	1
11B3	*let-329*	1	1
14	*let-472*	1	1
10	*let-452*	1	1
11B2	*let-473*	1	1
20B	*let-454*	1	1
16	*let-474*	1	1
10	*let-331*	1	1
20B	*let-475*	1	1
20A	*let-467*	1	1
13	*let-476*	1	1
14	*let-469*	1	1
10	*let-477*	1	1
1C	*let-453*	1	1
11B3	*let-463*	1	1

Table 6 (cont)

ZONE	GENE	TOTAL NUMBER OF ALLELES	EMS ALLELES INDUCED OVER *eT1*
15	*let-468*	1	1
17	*let-456*	1	1
6A	*let-478*	1	1
11B2	*let-479*	1	1
16	*let-480*	1	1
10	*let-481*	1	1
7	*let-461*	1	1
11A1	*let-443*	1	1
17	*let-445*	1	1
8A2	*let-338*	2	1
10	*let-350*	2	1
14	*let-410*	2	1
5	*let-426*	2	1
1B1	*let-447*	2	1
16	*let-423*	2	1
4A	*unc-60*	3	1
16	*let-405*	3	1
12A	*let-402*	3	1
20B	*let-407*	4	1
4B	*emb-29*	4	1
10	*let-428*	2	2
9A	*let-344*	2	2
2	*let-431*	2	2
8A2	*let-455*	2	2
18	*let-412*	2	2
18	*let-464*	2	2
8A2	*let-466*	2	2
9A	*let-459*	2	2
19	*let-340*	2	2
19	*let-334*	2	2
4A	*let-326*	2	2
11B2	*let-429*	2	2
20B	*let-470*	2	2
10	*let-415*	3	2
11B1	*let-349*	3	2
12B	*let-403*	4	2
10	*let-417*	4	2

TABLE 6 (cont)

ZONE	GENE	TOTAL NUMBER OF ALLELES	EMS ALLELES INDUCED OVER *eT1*
15	*let-339*	3	3
10	*let-419*	3	3
17	*let-424*	3	3
1B1	*let-336*	3	3
12A	*let-444*	3	3
11B2	*let-462*	3	3
6A	*let-327*	3	3
15	*let-442*	3	3
9A	*let-348*	4	3
17	*let-413*	4	3
16	*let-411*	4	4
11B4	*let-421*	4	4
11B2	*let-439*	4	4
13	*let-337*	5	4
9B	*let-345*	5	4
11B5	*let-422*	7	4
9B	*let-342*	5	5
10	*let-440*	5	5
15	*let-343*	6	5
14	*unc-70*	7	5
19	*let-409*	7	5
16	*let-335*	6	6
15	*let-346*	7	6
10	*let-420*	7	6
8A1	*lin-40*	15	6
9B	*let-341*	8	7
15	*let-332*	8	8
18	*rol-3*	13	8
7	*let-330*	17	16

Figure 12

Genetic map of the *eT1* balanced region of LGV(left). Descriptions of the deficiencies' origins (except *sDf56*) may be found in 1) Rosenbluth *et al.* 1988; 2) Johnsen and Baillie 1988 or Section 1; 3) Clark *et al.* 1990 or Section 2; 4) Stewart *et al.* 1990.

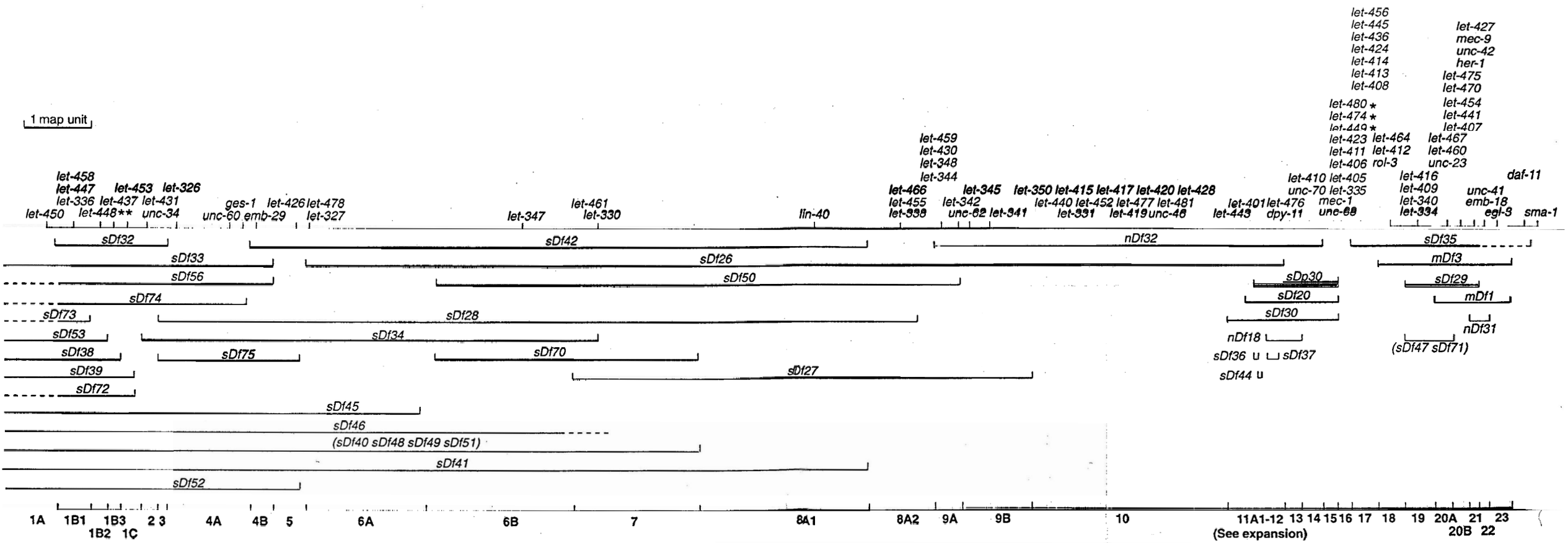
* *let-449*, *let-474* and *let-480* were placed in zone 16 because they complemented all other zones. They have not been three-factor mapped to the zone 16 region.

** The left end of *sDf74* could be within *sDf32* and then *let-448* could be to the left of *let-336*, *let-447* and *let-458*.

In the expansion of 11A1-15:

1 *let-349* and *let-422* could be in either zone 11B1 or 11B5, their locations on the map are based on two-factor mapping distances from *unc-46* (zone 10).

2 *sDf36* could overlap *sDf44* to the right, if so 11B1 would be to the right of 11B2 and 11B3 would be to the left of 11B2.


1 map unit
let-458
let-447
let-336
let-450
let-448**
let-437
let-453
let-431
unc-34
let-326
sDf32
ges-1
unc-60
emb-29
let-426
let-478
let-327
sDf42
let-347
let-461
let-330
lin-40
let-459
let-430
let-348
let-344
let-466
let-455
let-338
let-345
let-342
unc-62
let-350
let-415
let-417
let-420
let-428
let-341
let-440
let-452
let-477
let-481
let-331
let-419
unc-46
nDf32
let-401
let-443
let-476
dpy-11
let-410
unc-70
let-405
let-335
mec-1
unc-68
let-456
let-445
let-436
let-424
let-414
let-413
let-408
let-480 *
let-474 *
let-449 *
let-423
let-411
let-406
let-464
let-412
rol-3
let-427
mec-9
unc-42
her-1
let-475
let-470
let-454
let-441
let-407
let-467
let-460
unc-23
let-416
let-409
let-340
let-334
daf-11
unc-41
emb-18
egl-3
sma-1
sDf35
mDf3
mDf1
nDf31
sDf29
(sDf47 sDf71)
sDf33
sDf56
sDf26
sDf50
sDp30
sDf20
sDf30
nDf18
sDf36
sDf37
sDf44
sDf74
sDf73
sDf53
sDf38
sDf39
sDf72
sDf28
sDf34
sDf75
sDf70
sDf27
sDf45
sDf46
(sDf40 sDf48 sDf49 sDf51)
sDf41
sDf52
1A
1B1
1B3
1B2
1C
2
3
4A
4B
5
6A
6B
7
8A1
8A2
9A
9B
10
11A1-12
(See expansion)
13
14
15
16
17
18
19
20A
20B
21
22
23

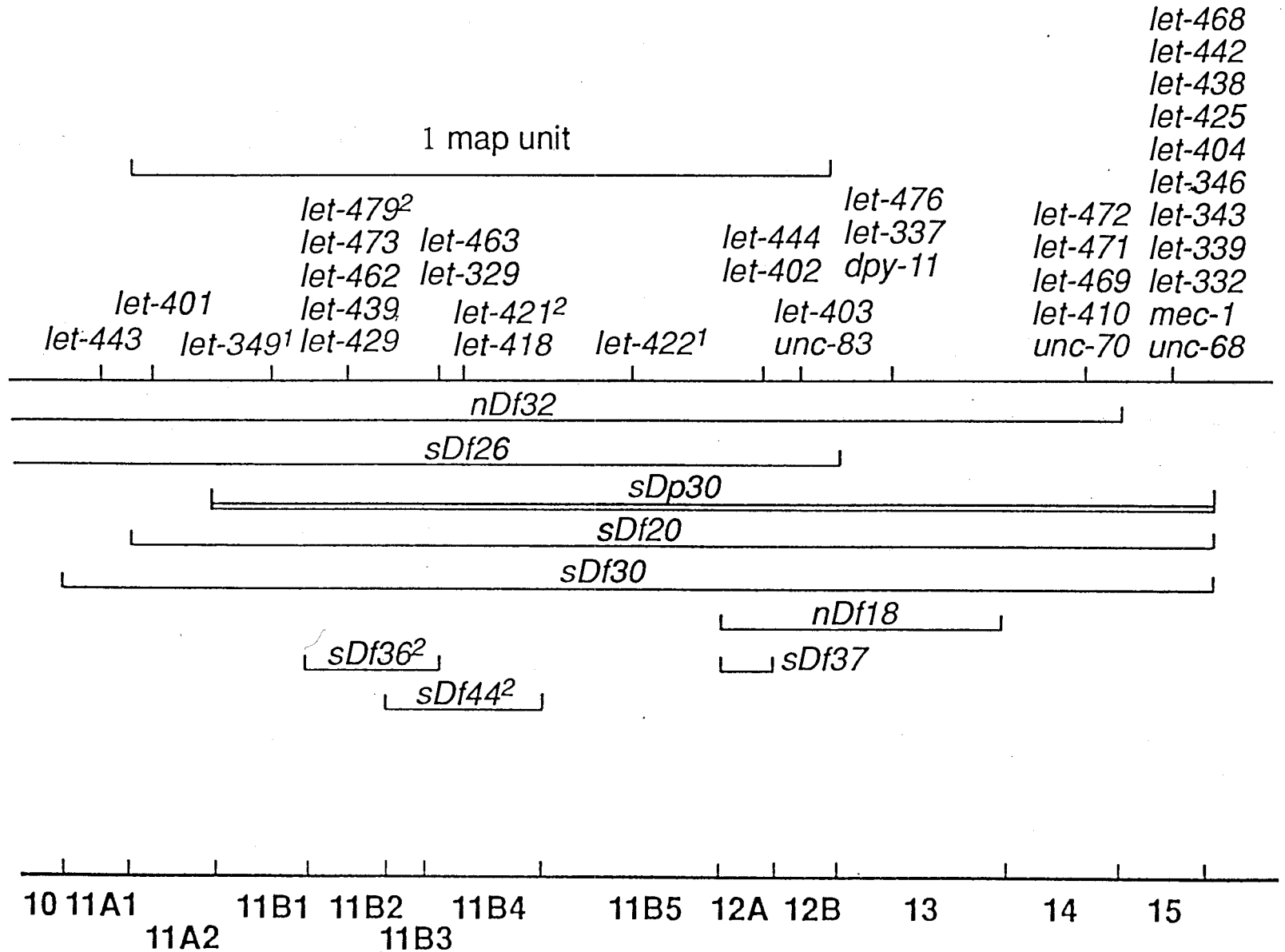

1 map unit
let-468
let-442
let-438
let-425
let-404
let-346
let-343
let-339
let-332
mec-1
unc-68
let-479²
let-473
let-462
let-439
let-429
let-463
let-329
let-421²
let-418
let-422¹
let-444
let-402
let-476
let-337
dpy-11
let-403
unc-83
let-472
let-471
let-469
let-410
unc-70
let-401
let-443
let-349¹
nDf32
sDf26
sDp30
sDf20
sDf30
nDf18
sDf37
sDf36²
sDf44²
10
11A1
11A2
11B1
11B2
11B3
11B4
11B5
12A
12B
13
14
15

Figure 13

Distribution of the number of alleles for the 89 genes on LGV(left) identified from the total EMS set. The expected values were calculated by the Poisson formula based on a mean number of alleles per gene of 1.25 (calculated from the truncated Poisson, see Discussion). *let-330* is the extremely mutable gene with 16 alleles.

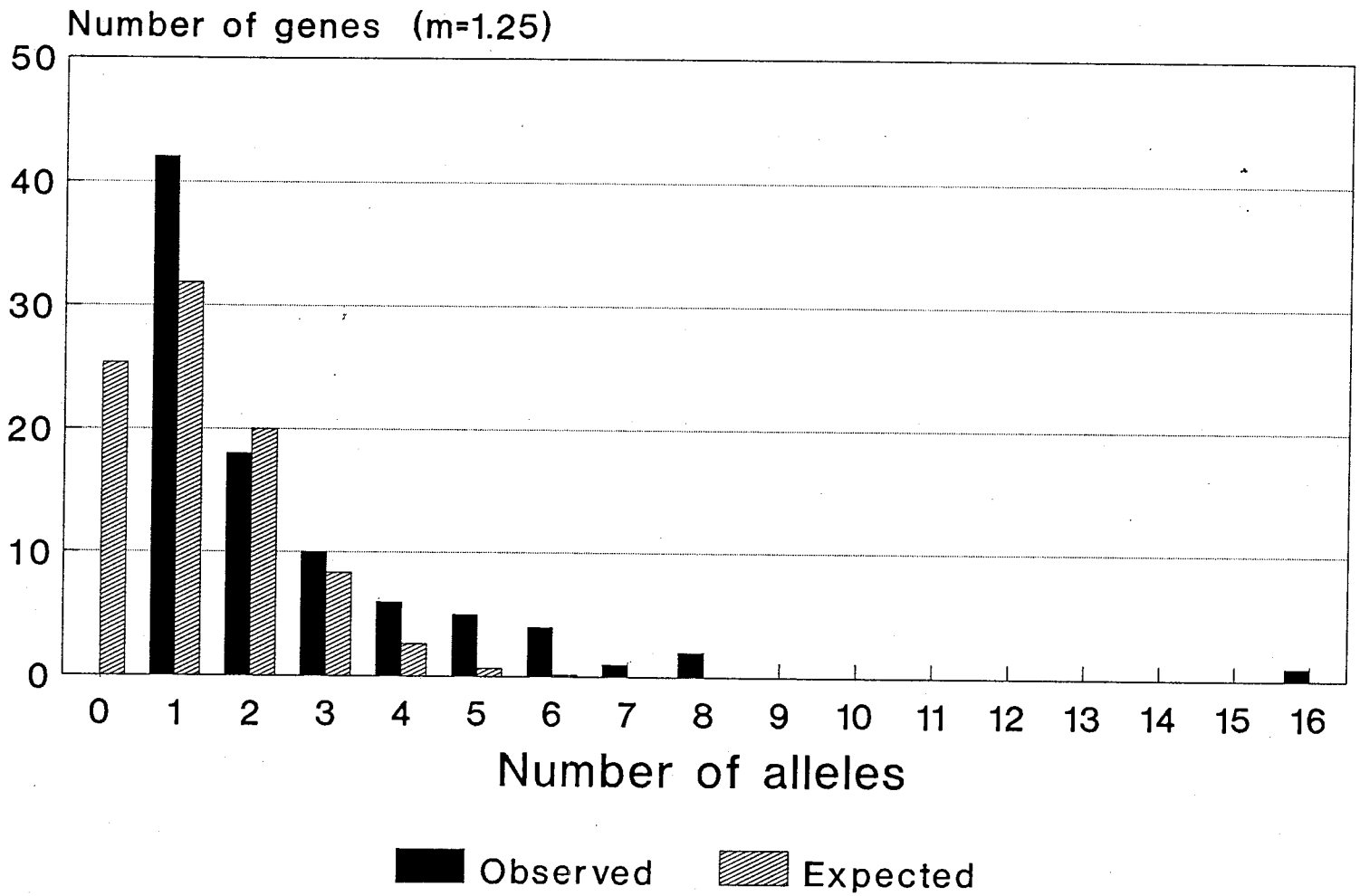

Number of genes (m=1.25)
50
40
30
20
10
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Number of alleles
Observed
Expected

Table 7
(Temperature rescue of EMS induced mutations)

Zone	Gene	Allele	15C	25C	Comments
1B1	*let-458*	*s1443*	no	N.D.	
4B	*emb-029*	*s1613*	no	N.D.	ts alleles exist
7	*let-461*	*s1486*	partial	no	sel and lag alleles
8A1	*lin-040*	*s1634*	no	N.D.	
8A2	*let-455*	*s1447*	no	N.D.	
10	*let-477*	*s1608*	no	N.D.	
11B4	*let-418*	*s1617*	partial	no	
11B5	*let-422*	*s1578*	yes	no	
13	*let-476*	*s1621*	yes	N.D.	male rescued
15	*let-343*	*s1465*	no	partial?	
15	*let-442*	*s1430*	partial	no	
15	*let-346*	*s1630*	no	N.D.	male rescued
18	*let-464*	*s1504*	no	N.D.	
19	*let-409*	*s1507*	no	N.D.	
20B	*let-470*	*s1629*	yes	no	male rescued
20B	*let-441*	*s1414*	yes	no	leaky at 20C
	let-?	*s1446*	no	partial?	stopped working on
	let-?	*s1633*	no	N.D.	slow stopped work on

Note: The strains that contained more than one lethal mutation were not tested for temperature rescue.

Figure 14

Plot of the number of genes whose most severe alleles block development at the indicated stages.

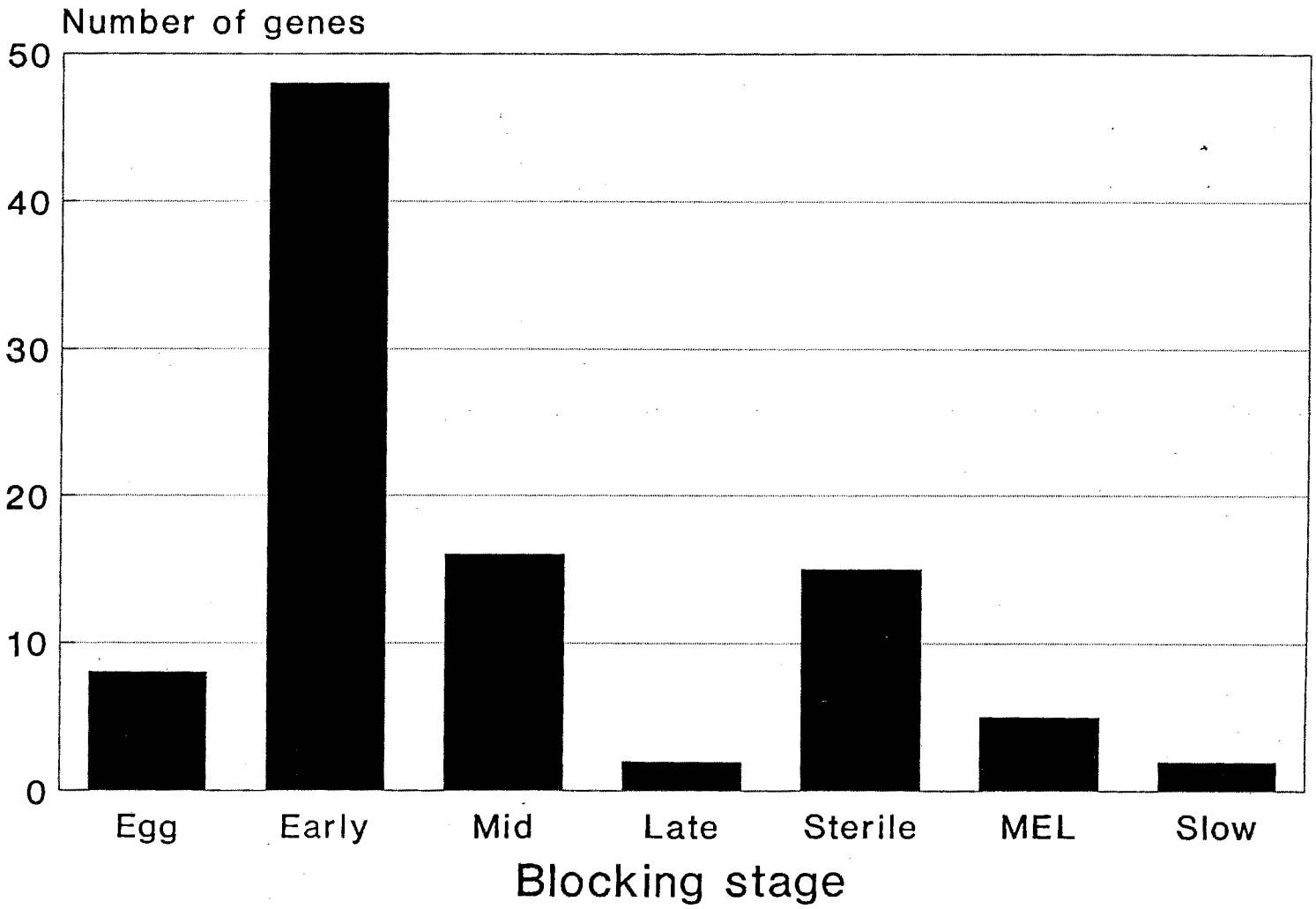

Number of genes
50
40
30
20
10
0
Egg
Early
Mid
Late
Sterile
MEL
Slow
Blocking stage

Table 8
(Male rescue of EMS induced mutations)

Zone	Genes	allele	Blocking Stage	Rescued
1B1	*let-447*	*s1457*	Maternal Mel egg lethal	no
2	*let-431*	*s1044*	Ster Adult	N.D.
6A	*let-327*	*s1496*	Slow	N.D.
8A1	*lin-040*	*s1611*	Ster Adult Vulva blip	no*
8A2	*let-466*	*s1063*	Maternal Mel Ster Adult	N.D.
9A	*let-348*	*s1448*	Maternal Slow Mel egg lethal	no
9A	*let-430*	*s1042*	Ster Adult	N.D.
9A	*let-459*	*s1615*	Maternal Mel egg lethal	no
10	*let-420*	*s1478*	Ster Adult Vulva blip	no**
10	*let-477*	*s1608*	Slow	N.D.
11B2	*let-429*	*s1597*	Slow Some morphological abnormalities	no
11B2	*let-479*	*s1576*	Ster Adult Lays mushy eggs	yes
11B4	*let-418*	*s1617*	Ster Adult Vulva blip (partial ts)	no
12A	*let-444*	*s1569*	Maternal Mel egg let. (not canonical)	yes
13	*let-476*	*s1621*	Ster Adult (ts)	yes
14	*let-469*	*s1582*	Ster Adult	yes
14	*let-471*	*s1570*	Maternal Mel larval lethal	yes
15	*let-346*	*s1630*	Ster Adult Not canonical allele	yes
15	*let-425*	*s385*	Ster Adult	N.D.
16	*let-480*	*s1607*	Ster Adult	no
17	*let-424*	*s1587*	Ster Adult	no
17	*let-445*	*s1419*	Maternal Tight coilers; Mel	N.D.
18	*let-412*	*s1598*	Ster Adult Roller	no
20B	*let-470*	*s1629*	Slow Not canonical allele	yes
20B	*let-475*	*s1606*	Ster Adult (ts)	yes
	let-?	*s1446*	Ster Adult Stopped working on	yes

* *lin-40(s1634)* another sterile adult allele of was not male rescued.

** Three other alleles of *let-420* (*s1573*, *s1584*, and *s1603*) were not male rescued.

Figure 15

Plots of the number of genes and the number of mutations mapped to genes versus the number of chromosomes screened after 0.012M EMS mutagenesis.

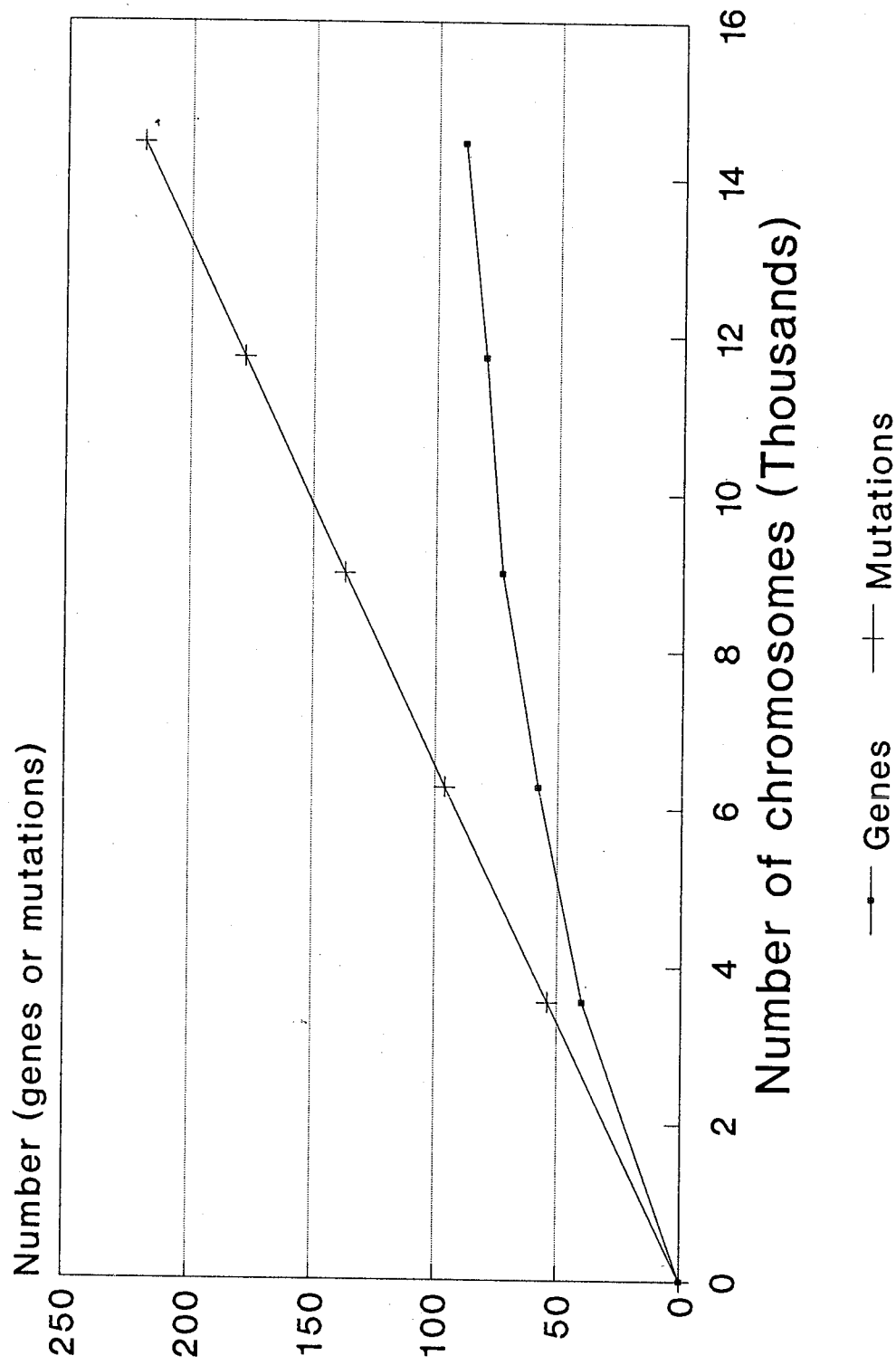
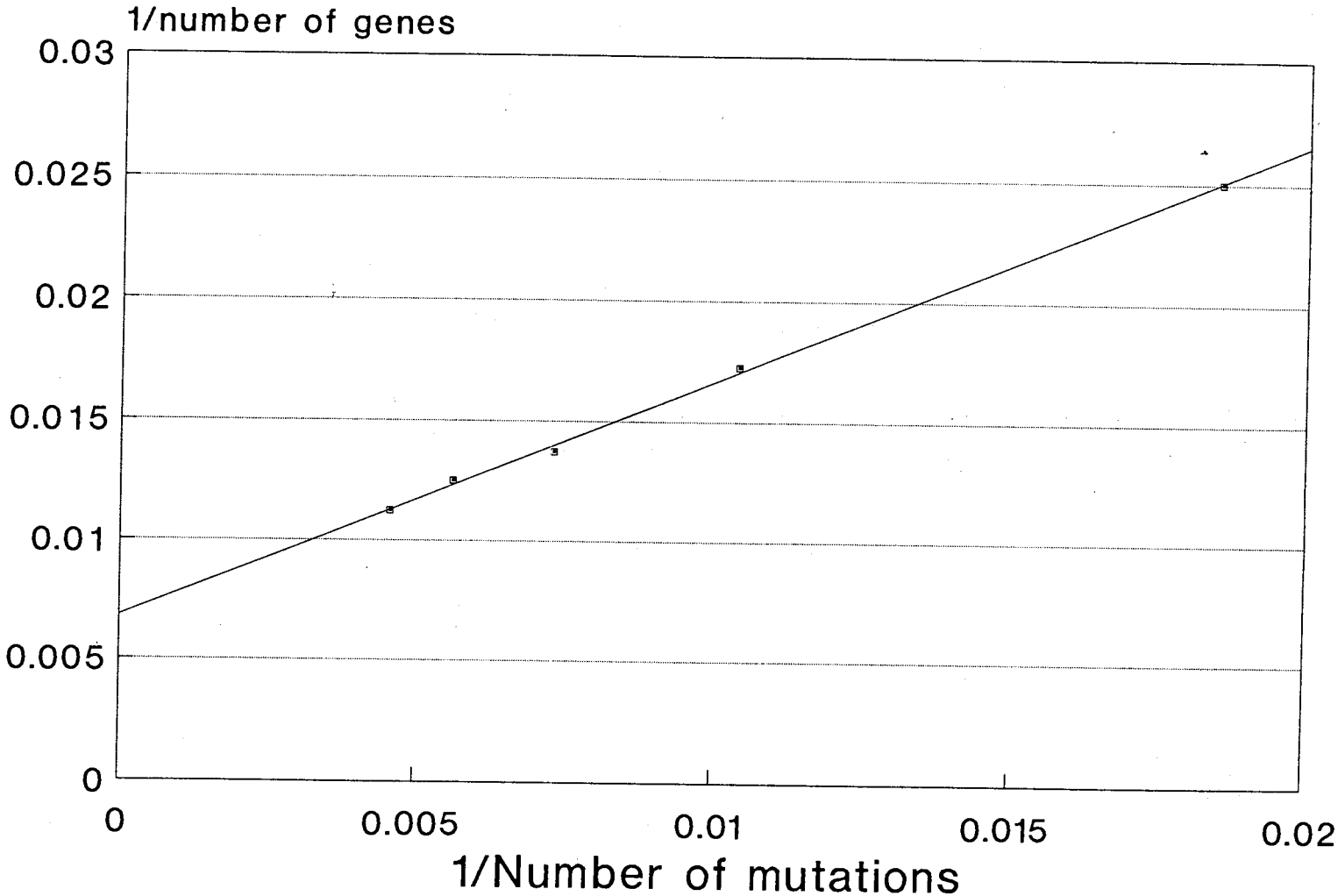

Number (genes or mutations)
250
200
150
100
50
0
0
2
4
6
8
10
12
14
16
Number of chromosomes (Thousands)
Genes
Mutations

Figure 16

Double reciprocal plot of the number of genes identified versus the number of mutations mapped to genes. The y-intercept is the point where an infinite number of mutations would have been analyzed and therefore the reciprocal of this value (147) is an estimate of the total essential gene number on LGV(left).


1/number of genes
0.03
0.025
0.02
0.015
0.01
0.005
0
0
0.005
0.01
0.015
0.02
1/Number of mutations

DISCUSSION

In a preliminary screen I reported that two-thirds of the lethal mutations induced by EMS in the *eT1* balanced region mapped to LGIII and one-third mapped to LGV (Johnsen *et al.* 1986). My results from the analysis of the 751 EMS induced lethal mutations isolated in this section supports this result: 64% of the mutations mapped to LGIII and 36% mapped to LGV. Since the two regions are recombinationally approximately equal in size, this implies that the gene density on LGIII(right) is twice that of LGV(left).

I have isolated 242 strains containing mutations that mapped to LGV(left), but 50 (27%) of the strains later proved to have their lethal mutations outside the balanced region. When mapping to a translocation balanced region, mutations outside the region but recombinationally close to it could appear balanced at first but would eventually cross away. *eT1* breaks in the main gene cluster of LGV and therefore there are a lot of essential genes just outside the balanced region which provides an explanation for the large number of non-balanced mutations recovered. The remaining 192 strains contained 194 lethals. I mapped 166 of the lethals and found them to be alleles of 78 genes and one deficiency.

To estimate the number of essential genes on LGV(left) and in the genome, I used two methods 1) a truncated Poisson calculation and 2) a double reciprocal plot of the number of genes identified versus the number of mutations analyzed. For the first method, a truncated Poisson is preferable to a normal Poisson calculation, but any Poisson is only useful for giving a minimum estimate of the number of genes. The reason for this is that the Poisson distribution requires the incorrect assumption that the probabilities of both inducing and recovering mutations for all genes is the same. The mutability of genes varies considerably (Meneely and Herman 1979; Hilliker *et al.* 1981; Lefevre and Watkins

1986; Howell 1989; Clark 1990). The genes that are large targets for mutagenesis are over represented and the small targets are under represented. Both biases lead to an underestimation of the gene number. A truncated Poisson lessens the effect of the over represented genes. m (mean number of mutations per essential gene) was calculated using f (the ratio of genes represented by more then one allele to genes represented by one allele): $f=[1-(e^{-m}-me^{-m})/me^{-m}]$. From f and a graph of m versus f, m could be determined. Figure 13 shows the distribution of allele frequency compared to that predicted by the Poisson distribution. The allele distribution is clearly not the same as the Poisson distribution. For the second method, I looked at the increase in the number of known genes with respect to the number of mutations analyzed. To do this, I divided the total EMS set into five groups, analyzed them sequentially and plotted the number of genes and the number mutation versus number of chromosomes screened (Figure 15). The first of the five groups from the total EMS set was the 54 mutations from Rosenbluth *et al.* (1988). Since these had been recovered after treatment with 0.004 - 0.0058M EMS, the expected number of 0.012M EMS treated chromosomes needed to be screened to identify 54 mutations on LGV(left) (that can be mapped without unreasonable effort) is 3545. For the other four groups I divided my EMS set (I did not include *sDf56*) into four random groupings of 42, 41, 41 and 41 mutations and noted the number of genes identified by each group. I then plotted the reciprocals of the number of genes versus number of mutations (Figure 16). In Figure 16 I did a best fit line (Harvard Graphics Type "trend"), and noted where the line crossed the y-axis. This represented the point where the number of mutations analyzed goes to infinity and therefore the reciprocal of the intercept value is an estimate of the total number of genes on LGV(left). The value noted was 0.068 the reciprocal of which is 147 genes.

There are two main problems with the method used for the double reciprocal plot: 1) The analysis of early screens would be biased by the large targets for EMS

mutagenesis. This would decrease the slope in Figure 16 and thus lead to an underestimate of the total gene number. 2) The points are cumulative, *i.e.* each succeeding point on the graph depends on the mutations and genes, used to generate the preceeding points. The way to eliminate this problem is prohibitive, it is to do a series of various sized screens and plot the results from each screen as an independent point. Despite these problems, I feel that this method is superior to the truncated Poisson calculation for deriving an estimate of the total gene number.

To estimate the number of essential genes in *C. elegans*' genome from the number of essential genes identified on LGV(left), one must know what fraction of the total number of genes are on LGV(left). Edgley and Riddle (1990) list 830 genes in *C. elegans*' genome, 244 are *lets* (essential genes) and most come from screens of specific regions (for LGI, see Rose and Baillie 1980; Howell *et al.* 1987; Howell 1989; LGII Sigurdson *et al.* 1984; LGIV Rogalski *et al.* 1982; Rogalski and Baillie 1985; Clark *et al.* 1988; Rogalski and Riddle 1988; LGV Rosenbluth *et al.* 1988; Johnsen and Baillie 1988; Clark *et al.* 1990; LGX Meneely and Herman 1979; 1981). Ignoring the *lets* there are 586 genes, 77 of which are on LGV (see genetic map of Edgley and Riddle 1990) . Assuming that there are no biases in the distribution, and that the essential genes follow the same distribution, then 13.1% of all *C. elegans*' genes are on LGV. 21 of the 77 genes on LGV are on LGV(left), 45 are on LGV(right), and 11 have not yet been mapped to position. Ignoring the 11 genes not positioned, 32.2% of the genes on LGV are on LGV(left) and therefore LGV(left) contains 4.2% of the essential genes in the genome. A total of 89 essential genes have been identified on LGV(left) (using the total EMS set) and assuming that all genes had an equal probability of being identified, then from the number of essential genes with one allele and the number with more than one allele, the truncated Poisson calculation yielded an average hit frequency per gene of 1.25 and a minimum estimate of 125 essential genes on LGV(left). From this it was calculated that are a minimum of 3000 essential genes in *C.*

elegans' genome. When I use the value from the double reciprocal plot the gene estimate on LGV(left) is 147 and the estimate for the total genome is 3500. Brenner's (1974) original estimate was 2000, Howell (1989) estimated 4000 and Clark (1990) estimated 3500. My estimate is similar to the latter two estimates.

Rosenbluth *et al.* (1983) presented a dose-response curve for EMS mutagenesis. Using this curve, I have adjusted the forward mutation rates cited by Brenner (1974), Moerman and Baillie (1981), Howell (1989) and Clark (1990) to the rates expected for 0.012M EMS mutagenesis in *C. elegans*. Brenner (1974) estimated the forward mutation rate for genes in *C. elegans* with visible phenotypes. His estimate becomes $2.4X10^{-4}$, adjusted for 0.012M EMS mutagenesis. Howell and Rose (1990) noted this estimate was based on a small sample size and a low estimate of the total number of genes and argued that the estimate was too high. Howell and Rose (1990) estimated the average forward mutation rate per essential gene in the *hDf6* region of LGI. Their estimate becomes $4.3X10^{-5}$ when I adjusted for 0.012M EMS mutagenesis. Clark (1990) estimated the rate for the *sDf2* region of LGIV. It is $5.8X10^{-5}$ adjusted for 0.012M EMS. Both Clark (1990) and Howell and Rose (1990) removed the highly mutable genes (discussed later) from their samples to calculate their forward mutation rates. My rate (after removing the highly mutable genes) for the total EMS set is $6.7X10^{-5}$ mutations per gene per chromosome screened for LGV(left). This is about one and one half times Howell and Rose's (1990) rate but is similar to Clark's (1990) rate. Although this rate is higher than Howell's, it is significantly lower (3 to 4 fold) than Brenner's (1974).

In order to identify 95% of the essential genes on LGV(left), enough chromosomes have to be screened to allow the recovery of 3 alleles for the average gene (from the Poisson distribution). For a forward mutation rate of $6.7X10^{-5}$, a minimum of 45,000 0.012M EMS treated chromosomes must be screened.

Five of the 96 essential genes on LGV(left) whose approximate blocking stages are known, have canonical alleles recessive for maternal effect lethality (*let-445*, *let-447*, *let-459*, *let-466* and *let-471*). Three were tested for male rescue: *let-447* and *let-459* were not rescued but *let-471* was. Therefore, at least two genes appeared to be strict maternal effect genes. From this, I calculate that there are at least 73 strict maternal effect genes in *C. elegans*' genome. Kemphues *et al.* (1988) identified 17 strict maternal effect genes which they divided into two classes: 13 genes that had one-two alleles and four genes that had four or more alleles. Because the smaller class had a forward mutation rate comparable to Brenner's (1974) rate, they considered the smaller class to be the real strict maternal genes. This led to a very low estimate (about 12) for the number of strict maternal genes in *C. elegans*' genome.

Howell (1989) described three classes of essential genes with respect to EMS induced mutability. The classes are: genes with average mutability; genes with much higher mutability; and rare extremely mutable genes. At least 13 genes on LGV(left) have high mutability (five or more alleles from the total EMS set) and one gene, *let-330* (Figure 13), is an extremely mutable gene (forward mutation rate of $1.1X10^{-3}$ with 0.012M EMS mutagenesis). This phenomenon of one (or a few) essential gene being of extreme mutability occurred in other chromosomal regions in which attempts were made to saturate for essential genes. Examples include *let-354* in the *hDf6* region of LGI in *C. elegans* (Howell 1989), which has a forward mutation rate of $6X10^{-4}$ at 0.012M EMS (15 times higher than her average hit frequency); *let-56* and *let-65* in the *sDf2* region of LGIV in *C. elegans* have forward mutation rates of $4X10^{-4}$ at 0.012M EMS (Clark 1990). *Picolo* is an example in the *rosy* region in *D. melanogaster*, where 153 mutations were mapped to 20 complementation groups and *Picolo* was hit 20% of the time, which is nearly twice as often as the next most hit gene (Hilliker *et al.* 1980). In the case of *Zw1* on the X-chromosome of *D. melanogaster*, putative point mutations were recovered in 16 genes in

the 3A2-3C2 region, nearly 25% of the mutations were in *Zw1* (Judd *et al.* 1972). These findings imply a discontinuity in the sizes of genes: there is a set of genes an order of magnitude larger than average. There is probably also a class of small genes (such as the the tRNAs). These genes would be underrepresented in EMS mutagenesis screens.

Figure 14 shows the number of genes whose canonical alleles block development at the indicated stages. If one assumes that LGV(left) is representative of the genome and that the genes shown in Figure 14 reflect the proportion of genes needed at the different developmental stages then the proportion and minimum number of genes needed at each stage can be calculated. From these assumptions, wild-type genes needed in the zygote for normal egg development consist of 8% (300) of all the essential genes. To get beyond early larval development requires 50% (1750) of the essential genes. To get past the mid-larval stage requires 17% (580) of the essential genes. To get past late larval development requires 2% (73) of the essential genes. 16% (550) of the essential genes are needed to make fertilized eggs. These include the genes needed for the formation and development of viable oocytes and viable sperm. My estimate for the number of genes necessary for the formation and development of viable sperm is 75 (discussed later). A further 5% (182) of the essential genes are needed make viable progeny, these are the *mel* genes of which at least 73 are strictly maternal. It is interesting that half of the essential genes appear to be needed strictly for early larval development. This may be a misleading result for these two reasons: 1) The products of many of the genes may be required much earlier but are maternally supplied and 2) the early larval blocking stage could be due to the metabolic demands of moulting *i.e.* sickly worms could live for awhile but not make it through a moult, even if all the genes needed strictly for moulting were wild-type. The decreasing number of genes necessary to develop through mid and late larval stages seems reasonable and they are probably involved in the specifics of the later moults. Although some worms may have mutations unrelated to moulting stage, they may be dying worms that make it

through one or two moults. The relatively large number of genes required to get past the mid-larval stage could be related to the fact that a developmental choice is made there. Under conditions of starvation and overcrowding *C. elegans* nematodes can enter a dispersive larval stage called dauer. Dauer-constitutive mutations may cause the worms to get locked into the dauer development pathway. If that happened the worms would not grow longer than the length that I assigned as mid-larval and thus be called mid-larval lethals. The large number of genes needed strictly to make fertile animals probably reflects the complexity of the gonad.

Male rescue is not only useful for identifying strict *mel* genes but also for identifying mutations that cause defective spermatogenesis (Spe) or defective fertilization (Fer) [*fer* is no longer used as a designation for new genes in these categories (discussed by Kimble and Ward 1988)]. Mutations in *spe* genes are ones that result in hermaphroditic worms becoming sterile adults. From preliminary results I have identified eight mutations that have been mapped to genes on LGV(left) that are rescued by N2 sperm (Table 8). I categorized the mutations according to their effects on homozygote hermaphrodites: one is a *mel* larval lethal that is male rescued [*let-471(s1570)*]; two are weak alleles of genes whose canonical alleles block development at an early stage [*let-444(s1569)*] or a mid-larval stage [*let-346(s1630)*]; three are both male rescuable and temperature sensitive [*let-470(s1629)*, *let-475(s1606)* and *let-476(s1621)*]. All three genes are represented by only one allele. *s1629* is also a slow developer at 20C; two are male rescuable sterile adults: *let-479(s1576)* appears to produce oocytes continuously while *let-469(s1582)* does not.

The last two (*let-479* and *let-469*) are good candidates to be *spe* (defective spermatogenesis) genes. It should be noted that sterile adult male rescuable mutations can occur in genes whose canonical alleles are not sterile adults, an example from my EMS set is *let-346(s1630)*. At least 36 *spe* genes have been recovered (L'Hernault *et al.* 1988 and sources cited therein) in *C. elegans*' genome. L'Hernault *et al.* (1988) used a truncated

Poisson equation to calculate that there are a mimimum of 14 *spe* genes on LGI (42 m.u.). If LGI is representative of the 300 m.u. in the genome, then their estimate would be a minimum of 100 *spe* genes in the genome. My estimate (based on two genes) of a minimum of 73 *spe* genes is very similar to their estimate. From this it appears that *spe* genes are randomly located in the *C. elegans* genome.

Temperature sensitive mutations often result from a single amino acid substitution that renders a polypeptide inactive at some temperatures but not at others (Suzuki *et al.* 1967 and references cited therein). I tested 190 of my 194 EMS induced mutations for rescue at 15C [the four not tested were in the two strains that had two mutations each on LGV(left)]. Four (2.1%) of the mutations were rescued and three others were partially rescued (Table 7). Therefore seven (3.7%) were at least partially ts. Suzuki *et al.* (1967) found a similar level (6.3%) of all EMS induced lethals on the X chromosome in *D. melanogaster* were ts. Because the P_0 "B" brood was maintained at 15C, some of the ts mutations in the F_1s from that brood might have been temperature rescued and therefore not isolated. I believe that this was not a major problem because two of the four ts rescued mutations came from the B brood (approximately one half of all the mutations in my EMS set came from the B brood). None of the 177 mutations tested for rescue at 25C survived although two were partially rescued. The four ts mutations that I had isolated were *let-422(s1578)*, *let-476(s1621)*, *let-470(s1629)* and *let-441(s1414)*. Other genes on LGV(left) have ts alleles, they include *emb-29* and *rol-3*. *let-327* has a cold sensitive dominant allele. Ts mutations are useful for a variety of analyzes including temperature shifts experiments to determine the temperature sensitive period (putative time of gene action), reversion experiments and suppression experiments to help determine what other genes interact with with the gene containing the ts allele.

The *unc* (uncoordinated) genes generally are involved in muscle or in the nervous system controlling muscles. The null phenotype of most identified muscle genes is Unc

(this could be due to the fact that the Unc phenotype is the main one used to identify muscle genes), but for a few genes it is early larval lethality. Some dominant missense lethals in *unc-54* are recessive lethals (MacLeod *et al.* 1977; Dibb *et al.* 1985; Bejsovec and Anderson 1988). *let-75* and *myo-3* are the same gene (pers. comm. A. Rose) and the gene has some early larval lethal recessive alleles (Rose and Baillie 1980). Waterston (1989) showed that the minor myosin heavy chain *(mhcA)* also has recessive lethal mutations. Waterston demostrated that worms homozygous for such mutations do not progress beyond 2 folds in the egg (as the worm elongates in the egg it must fold to accommodate its length, wild-types progress beyond 3 folds). The worms hatched but remained in a folded state with no locomotion. I have identified the first recessive lethal mutation of *unc-60* and five recessive lethal mutations of *unc-70* (one other recessive lethal allele of *unc-70* was isolated by Rosenbluth *et al.* 1988). The *unc-70* recessive lethal mutations caused a similar phenotype to that exhibited by *mhcA*. Although the state of folding in the egg was not noted, the eggs hatched and the worms persisted for several days curled or folded with no obvious growth and no sign of locomotion. Worms homozygous for the lethal allele of *unc-60* (*s1586*) also blocked development as early larval lethals but grew about one third longer than *unc-70* lethal homozygotes.

The genetic analysis of the large region [LGV(left)] has led to a number of interesting observations about genes and genomic organization. LGV(left) consists of two clusters, the main cluster (zones 8A2-20+) and a smaller cluster (zones 1-6A) near the end of the chromosome. *let-347-let-461-let-330-lin-40* reside in the high recombination/low gene density region between the two clusters (Figure 12). I will discuss this latter region first.

let-347: *let-347* is represented by one EMS induced allele (*s1035*) which I identified in a preliminary screen (Rosenbluth *et al.* 1988). *s1035* is recessive for late

larval lethality. *let-347* has been separated from *let-330* and *let-461* by the left breakpoint of *sDf27*.

let-461: This gene was called *lag-2* by J. Kimble and *sel-3* by J. Thomas. *let-461* has alleles that have the same phenotype as *lin-12 glp-1* double mutants (J. Kimble and E. Lambie pers. comm.) and alleles that are semi-dominant suppressors of *lin-12* dominant alleles (J. Thomas pers. comm.). One allele, *(s1486)*, is a recessive EMS induced mutation and results in the lag (*lin-12 glp-1*) phenotype at 20C but is partially rescued at 15C where it causes mid larval lethality. This suggests that *s1486* is not a null.

let-330: *let-330* is a husge gene with respect to EMS mutagenesis, with a 0.012M EMS forward mutation rate of $1.1x10^{-3}$. This is greater than the rate of $6.4X10^{-4}$ for *unc-22* (unpublished results cited in Moerman and Baillie 1981). *unc-22* spans 50Kb and encodes the muscle protein twitchin which has some homology to the vertebrate striated muscle protein titin (Benian *et al.* 1989). Twitchin has a molecular weight of approximately 700,000 daltons (Benian *et al.* 1989). From the map (Figure 12), one can see that *let-461* has not been separated from *let-330*. *sDf34* comes in from the left and deletes *let-330* and *let-461* but no other known genes to the right. The right breakpoints of *sDf34* and four deficiencies induced by the mobilization of Tc1 (*sDfs 40,41,48,49* and *51*) have not been separated. *sDf27* comes in from the right and deletes *let-330* and *let-461* but no other genes to the left. *sDfs 26,42* and *50* delete *let-330* and genes to both sides. Two alleles of *let-330* (*s1497* and *s1518*) behave differently over *sDf34* than over *sDfs 26* or *42*. Worms heterozygous for either *s1497* or *s1518* over *let-330(s573)*, *sDf26* or *sDf42* block as early-mid larvae, but over *sDf34*, mature to adulthood and give occasional larval progeny. *let-330(s1638)/let-330(s1433)* is a Mel larval lethal (implying possible interallelic complementation) but *let-330(s1638)/sDfs 27,34* or *42* are larval lethal. These data imply that: *let-330* is a large gene; *sDf34* breaks within *let-330*; *let-330* probably transcribes right-left; *let-461* is either internal to *let-330* or to its left; and *let-330* is a multi-domain

gene and/or encodes for proteins that function as multimers. If *sDf34* does break within *let-330*, then the molecular identification of the right breakpoint of *sDf34* could lead to the cloning of *let-330*.

lin-40: The developmental blocking stages of the 16 recessive *lin-40* alleles range from early-mid larval lethality to maternal effect early larval lethality (see Section 5). Worms homozygous for eight of the alleles develop into adults (sterile or maternal effect lethal) with protruding vulvae. The results of interallelic complementation tests show that *lin-40* is a complex locus with at least five complementation classes. The blocking stages of the various alleles of *lin-40* appear to vary consistently with their complementation classes. Some alleles of *lin-40* and deficiencies that delete *lin-40*, dominantly cause an increase (approximately 40%) in the number of self-progeny from *C. elegans* hermaphrodites (similar effects on brood size were observed for *let-336* in zone 1B1). This implies that *lin-40* not only affects the vulva-lineage but also affects the switch from spermatogenesis to oogenesis in hermaphrodites.

The *let-347-let-461-let-330-lin-40* region contains genes with interesting characteristics, is highly subject to being mutated by the mobilization of Tc1 (Section 2) and is a region of very high recombination. There are approximately six m.u. with four known genes in a region that is ~70% saturated, therefore, assuming approximately equal spacing of genes along the chromosome, the recombination rate is greater than 10 times the genome average. By looking at Figure 12, one can see that the deficiencies that are in the clusters are shorter (in map units) than the deficiencies in the gene sparce region. If two assumptions are made: 1) gamma irradiation and formaldehyde induced breakpoints occur at random; 2) there is not excessive haplo-insuffiency involved in deleting a large number of genes in the clusters (limiting the deficiency size); then the recombination rate in the *let-347-let-461-let-330-lin-40* region is much higher then average. Two factor distances for the genes in that region and *unc-46* are: *let-347(s1035)* 14.8 (12.8-16.8)

m.u., *let-330(s573)* 8.3(5.7-10.9) (Rosenbluth *et al.* 1988) and *lin-40(s1352)* 4.7(3.4-6.2) (Section 5). The multi-domain aspects of *lin-40* and possibly *let-330* are reminiscent of the bithorax and antennapedia complexes in *Drosophila.* Intergenic and fine structure mapping of the alleles of these genes, and the molecular analysis of the DNA sequence responsible for the high recombination, would help in elucidating the molecular basis for the high level of recombination. It is possibly due to a single site and therefore the interallelic and possibly the intergenic recombination of these genes would not be high. But if the interallelic recombination is high, then we must consider that the ability to reshuffle the domains of alleles of these genes could be evolutionarily important. An example of a system that uses interallelic recombination to generate diversity is the mammalian immune system. The natural population of *C. elegans* should also be analyzed for polymorphisms of these genes which could be fodder for selection. Further analysis of these genes, and genes in other regions of high recombination, should lead to insights into the nature of the non-random recombinational distribution of genes first noted by Brenner (1974). The DNA sequence responsible for the high recombination could also prove valuable as a recombinant molecular genetics tool.

***let-447* (zone 1B1):** *let-447* has two alleles (*s1457* and *s1654*). *s1654* was isolated from a screen of formaldehyde treated chromosomes (Johnsen and Baillie 1988 or Section 1). Both alleles show a recessive maternal effect egg lethal phenotype. Generally, homozygous *let-447(1457)* and *let-447(s1654)* hermaphrodites lay very few eggs, but both *s1447/Df* and *s1654/Df* hermaphrodites consistently lay many eggs. The blocking stages of the eggs in all four cases (homozygous or hemizygous *s1457* or *s1654*) appear to be the same: a multicellular mass of fewer then 550 cells [they did not reach the lima bean (start of morphogenesis) stage, see von Ehrenstein and Schierenberg 1980; Sulston *et al.* 1983]. I can propose two possible explanations for these results: 1) *let-447* has two functions: a) it is involved with the production of germ cells; and b) it has a maternal effect on egg

development. Both alleles of *let-447* would be hypermorphic for the first function and probably null for the second. An additional complication is that both alleles are recessive and show no obvious effect over the wild-type allele. 2) The increased number of eggs is not due to *let-447* but due to the deficiency. *let-336* is deleted by all the currently known deficiencies that delete *let-447*. Heterozygous *let-336(s1413)* or *let-336* (wild-type) over a deficiency cause increased brood sizes (Section 5). Thus, the increased egg production noted for *let-447* over a deficiency is really due to the hemizygosity of the neighbouring gene *let-336*. A deficiency that deletes *let-447* but not *let-336* would help in chosing between the two explanations.

***let-342* (zone 9B):** There are five alleles of *let-342*, all of which are EMS induced and from the total EMS set, this implies that *let-342* is a large target for EMS induced mutagenesis. Three of the alleles block development at the early larval stage (*s1442*, *s1487* and *s1549*), the other two (*s1029* and *s1616*) block at the mid larval stage. Occasionally worms homozygous for *s1616* mature and have internal larval progeny. I have placed *let-342* into zone 9B. All of the alleles over *sDf27*, *nDf32* or *let-342(s1029)* appear to block at early (or mid) larval stages (*s1487/s1549* blocks early, the other interallelic complementation tests have not been done). *s1442* complements *sDf50*, but worms heterozygous for any of the other alleles and *sDf50* block as sterile adults. Worms heterozygous for either *s1487* or *s1549* over *sDf50* were rescued by male sperm (*s1029* and *s1616* were not tested for male rescue). To explain these data I have two proposals 1) *sDf50* breaks near the promoter region of *let-342* and modifies its level of transcription or 2) *let-342* is transcribed right to left and *sDf50* breaks within *let-342* near the 3 prime end. In both proposals the chromosome containing *sDf50* produces almost normal levels of *let-342* product or almost normal product. This allows worms containing the *sDf50* chromosome over a putative null *let-342* allele, to mature and become male rescuable

sterile adults. If *sDf50* does break very close to or in *let-342* then by molecularly identifying the right breakpoint of *sDf50*, one should be able to clone *let-342*.

***let-341* (zone9B):** *let-341* is in the same zone as *let-342* and, to date, is the fourth biggest target for EMS mutagenesis on LGV(left). The canonical allele of *let-341* (*s1031*), which I isolated (Rosenbluth *et al.* 1988) is a recessive early larval lethal. Some *let-341* alleles have been shown to be suppressors of multivulva *lin-15(X)* mutations (S.G. Clark and H.R. Horvitz pers. comm.). *lin-34/let-60* suppresses the lethality of some alleles of *let-341* including *s1031* (S.G. Clark and H.R. Horvitz pers. comm.).

***let-420* (zone 10):** Worms homozygous for any of the seven alleles of *let-420* are sterile adults with protruding vulvae. *let-420* appears to be a large target for EMS mutagenesis. Six of the alleles were induced by EMS (one allele was induced by gamma irradiation), this induction frequency is over four times higher than the average hit frequency for essential genes on LGV(left).

***let-418* (zone 11B4):** Both alleles of *let-418* are recessive and show temperature sensitive effects. The terminal phenotypes of worms homozygous for *let-418(s1045)* include mid larval lethality at 15C, maternal effect early larval lethality at 20C and at 24C the adults are sterile and have protruding vulvae (R.E. Rosenbluth pers. comm.). Worms homozygous for *let-418(s1617)* are maternal effect lethals giving a few larval progeny at 15C but are sterile adults with protruding vulvae at 20C.

As described above, three of the identified essential genes (*let-418*, *let-420* and *lin-40*) have recessive alleles that not only cause adult worms to be sterile but also to have protruding vulvae and thus these genes putatively affect the vulva-lineage.

***rol-3* (zone 18):** I have identified five EMS induced recessive lethal alleles of *rol-3* from my EMS set (*s1408*, *s1409*, *s1473*, *s1494* and *s1519*), this increases the number of recessive lethal alleles of *rol-3* to 12 (the other seven were identified by Rosenbluth *et al.* (1988). All of the alleles from my EMS set (with the possible exception of *s1473*) are

recessive for an early larval lethal phenotype. *s1473* is in a strain with another recessive lethal mutation [*let-463(s2168)*]. The blocking stages due to these mutations individually were not determined. One of the *rol-3* alleles, which I identified in an earlier screen (see Rosenbluth *et al.* 1988), is temperature sensitive. At 20C it is a recessive lethal but at 15C it is a fertile roller. It was possible that the original roller phenotype *(e754)* was due to a spurious event [*i.e.* it was a neomorph or an antimorph (in the sense that it might have interfered with a normally unrelated gene product)]. The existence of the temperature sensitive allele that causes rolling at the permissive temperature allows me to argue that weak alleles cause rolling but stronger alleles cause lethality. Interestingly *let-412*, which has not been separated from *rol-3* by a chromosomal rearrangement, has an allele (*s1598*) which is recessive for a sterile adult roller phenotype. *s1598* has only been complementation tested against two *rol-3* alleles, *s833* and the "visible" allele *e754*, and complemented both. The possibility that *rol-3 - let-412* is a complex locus is being examined by W.B. Barbazuk. Another possibility is that *rol-3* and *let-412* are neighbouring genes that affect the same or related developmental pathway(s). The cloning and sequencing of these genes will resolve these alternatives.

Zone 16: This is a region that has not been deleted by any deficiencies. It is possible that there is a haplo-insufficient gene in the region which hinders the isolation of deficiencies. The location of most genes to the zone 16 region has been confirmed by 3-factor mapping (Rosenbluth *et al.* 1988), but 3-factor mapping has not been used to locate *let-449*, *let-474* or *let-480*. From 2-factor mapping data of *let-449* and *let-474* to *unc-46*, I got distances of 4.3 (2.1-9.1) and 4.8 (2.7-7.6) m.u. respectively. These distances are consistent with both genes being in zone 16. Two-factor mapping of *let-480* (an adult sterile) was not done. It is formally possible that one or more of these genes are to the left of all the deficiencies and reside near the left tip of the chromosome. As all three genes are represented by only one allele each, it is also possible that one or more of those alleles

are hypermorphic (generally rare alleles) such that worms heterozygous for the appropriate allele over a deficiency appear normal. If such is the case then the gene(s) could be anywhere on LGV(left).

The blocking stages listed in appendix B are generally based on the measured lengths of the worms. In many cases subtle distinctions between phenotypes would probably be missed. I made the assumption that the length of the worms indicated their stage of arrest. Generally, for genes with more than one allele, the blocking stages are approximately the same. This strongly suggests that most isolated mutations in genes are nulls. This is consistent with findings described by others (Shannon *et al.* 1972; Rose and Baillie 1980; Meneely and Herman 1981; Howell 1989; Clark 1990), although Clark (1990) found somewhat more variability. She often found that later arresting mutations blocked earlier as hemizygotes than homozygotes, indicating that they were hypomorphs (Muller 1932). In most cases I did not note the phenotypes of my mutations over deficiencies, but I did find that some anomalous late blocking alleles were temperature sensitive, thus possibly indicating that they were hypomorphs.

Because most mutations appear to cause null phenotypes, genes with several alleles with different phenotypes are good candidates for complex loci. An example is *lin-40* (Section 5), where the variation in blocking stages appeared to be consistent with the interallelic complemention map.

A large number of genes appear to have more than one seemingly unrelated phenotype. Some examples are: a) *let-336* which is a recessive early larval lethal but *let-336(s1413)* as a heterozygote causes increased brood sizes of hermaphrodites [presumably through delaying the switch from spermatogenesis to oogenesis (see Section 5)]; b) *lin-40* behaves similarly; c) and d) *let-461* and *let-341* have recessive early larval alleles, but they both have alleles that interact with gonad related lineage genes; e) *glp-1* is not only essential for the maintenance of germ cell/mitotic division by the distal tip cell in the gonad

(Kimble and White 1981) but is also required embryonically (Priess *et al.* 1987). Yochem and Greenwald (1989) showed that *glp-1* shares three repeated amino acids sequences with the *C. elegans' lin-12* and *D. melanogaster Notch* products. The three are: an epidermal growth factor-like motif; the "*lin-12/Notch* Repeat"; and a motif present in two yeast gene products that have cell cycle dependent functions. f) an example from Drosophila is *enable* (*ena*), a recessive embryonic lethal that has dominant alleles that compensate for the neural defects of *abelson* (*abl*) (Gertler *et al.* 1990). *abl* is a homolog of the mammalian *c-abl* cytoplasmic tyrosine kinase. A simple explanation for the fact that numerous genes have more than one time or place of action is that the genes are involved in general gene regulating systems such as: 1) DNA binding proteins that regulate gene expression and therefore the seemingly functionally unrelated genes that are acted on may have some of their promoter regions in common (other promoter regions must vary to allow differential control). 2) Signaling pathways; some of the genes may be ones that act in the parts (such as those involving the cytoplasmic tyrosine kinases) that are common to the regulation of otherwise unrelated genes. It is difficult to characterize genes that have more than one time or place of action, because one phenotype may interfere with the identification of the other phenotypes.

Researchers are working on the various developmental pathways of *C. elegans*. The identification of the majority of essential genes provides them with a collection of identified genes which can be quickly used to complementation test their putative developmental specific genes for possible essential alleles with other phenotypes.

Approximately 70% of the essential genes on LGV(left) have now been identified. It is the largest well characterized genomic region in any higher eukaryote. Hochman (1970) analyzed chromosome 4 in *D. melanogaster* and identified 33 essential loci and estimated that there were approximately 40 in total. Others have attempted to saturate regions in *D. melanogaster*, a recent example is the 17 polytene bands of 73A2-73B7 by

Belote *et al.* (1990). They identified 18 complementation groups including 13 that had not been previously identified. The complexity of most higher organisms makes the task of mapping all of their essential genes formidable. The ease of maintenance of *C. elegans* and the rapidity of its genetics makes it feasible to saturate the entire genome for essential genes. An attempt to do so is in the planning stages [D.L. Baillie, D.L. Riddle and A.M. Rose (pers. comm.)].

Section 5:

Two genes, *let-336* and *lin-40*, probably affect the switch from spermatogenesis to oogenesis in the hermaphrodite *Caenorhabditis elegans*

INTRODUCTION

The normal sexes of *Caenorhabditis elegans* are XX hermaphrodites and XO males. Hermaphrodites are considered to be somatically female but they are protandrous and produce approximately 320 sperm (see Hodgkin 1988) before switching to exclusive oocyte production. The ratio of X chromosomes to autosomes (X/A) is the primary signal for determining sex (Nigon 1949; Madl and Herman 1979); a ratio less than 0.67 results in male development and a ratio greater than 0.75 results in hermaphrodite development. An eight base pair non-coding sequence present in many copies on the X chromosome, appears to be the numerator in the X/A ratio (McCoubrey *et al.* 1988). The denominator has not yet been defined. The X/A ratio sets three known X-linked master regulator genes: *xol-1*, *sdc-1* and *sdc-2* (Miller *et al.* 1988; Villeneuve and Meyer 1987; Nusbaum and Meyer 1989). *xol-1* is the earliest known gene acting in the pathway controlling the male/hermaphrodite decision. Miller *et al.* (1988) proposed a model where *xol-1* negatively regulates the genes known to control both sex determination and dosage compensation (*sdc-1* and *sdc-2*).

Sex determination in *C. elegans* has been subdivided into three pathways (for a review see Hodgkin 1990): 1) Dosage compensation; 2) Somatic sexual development; and 3) Germ line sexual development.

Dosage compensation is needed to control the expression of X linked genes. Males with only one X chromosome have half the number of X chromosomes as hermaphrodites. For X-linked gene expression to be the same in both sexes, either expression in males must be increased and/or expression in hermaphrodites must be decreased. It is not yet known which occurs but failure to dosage compensate is lethal (Hodgkin 1983a; Meyer and Casson 1986; Meneely and Wood 1987; Delong *et al.* 1987; Miller *et al.* 1988).

Somatic sexual development refers to all sex specific differences in the somatic tissues. Hermaphrodites and males differ extensively in behavior and morphology (for a review of *C. elegans* morphology see White 1988). The earliest known developmental differences between the sexes occur during mid-embryogenesis and are the programmed cell deaths of four male specific sensory neurons in hermaphrodites and two hermaphrodite specific motorneurons in males (Sulston and Horvitz 1977; Sulston *et al.* 1983). The sex specific phenotypes are controlled by a pathway containing autosomal regulatory genes (Hodgkin 1987). The last gene in the pathway, *tra-1*, has been shown to act in a cell-autonomous manner (Hodgkin 1990 cites C. Hunter and W.B. Wood pers. comm.). *tra-1* is a global regulator affecting many developmental pathways (Hodgkin 1990). The phenotype, caused by the completely recessive null *tra-1* alleles, is the total somatic masculinization of XX animals, and generally, but not always, the complete masculinization of the gonad (Hodgkin and Brenner 1977; Hodgkin 1983b). Therefore, *tra-1*'s function is to regulate feminization morphology.

Germ-line sexual development refers to the type of gametes made by the gonads. In normal XX hermaphrodites a limited number of sperm are initially produced. Then, the gonads switch to exclusive oocyte production. Therefore, the germ line is initially male inside a female soma.

The pathway for germ-line sex is similar to the somatic sex pathway with most genes in common. The important differences are the activities of *fog-1* and *fog-2* in germ-line sex determination but not in somatic sex determination. *fog-2* is active early in hermaphrodites, and, acting through *tra-2*, results in the increased expression or activation of *fems-1,2,3* and *fog-1* which in turn induces spermatogenesis. Later, *fog-2* ceases to be active and *fems-1,2,3* and *fog-1* activity are decreased resulting in oogenesis (Doniach 1986; Schedl and Kimble 1988). Homozygous null alleles of both *fog-1* and *fog-2* have no effect on the soma of either hermaphrodites or males and both prevent spermatogenesis in

hermaphrodites allowing only oogenesis. *fog-1* causes oogenesis in XO somatic males but *fog-2* has no detectable effect on males (Barton and Kimble 1990).

To date the only candidate genes for controlling the switch from spermatogenesis to oogenesis are *tra-2* and *fem-3*. One weak allele of *tra-2*, *e1875*, has been shown to increase the mean number of self-progeny produced by an XX hermaphrodite (Hodgkin 1988). The null phenotype caused by *tra-2* is the recessive incomplete masculinizing effect on XX animals; the gonad is male with extensive spermatogenesis but the tail is abnormal and the animals do not show any mating behavior (Hodgkin and Brenner 1977; Hodgkin 1985). Gain-of-function *fem-3* alleles also cause the masculinization of the germ-line without affecting the soma (Barton *et al.* 1987).

In this Section, I discribe two genes on LGV(left) which have alleles that dominantly cause an increase in the mean number of self-progeny produced by XX hermaphrodites. Since the number of offspring produced by a hermaphrodite can be increased by mating to XO males, the number of self-progeny are apparently limited by the number of sperm produced during spermatogenesis. Thus the two genes appear to affect the switch from spermatogenesis to oogenesis. The genes are *let-336* and *lin-40*. Mutations in both genes have additional mutant phenotypes. All three alleles of *let-336* are recessive lethals, and the sixteen alleles of *lin-40* range from recessive lethals to recessive Mels. Because of the range of phenotypes caused by the *lin-40* alleles, the developmental abnormalities caused by various late blocking alleles were examined. All the abnormalities appear to be in sex-specific tissues. The results from *lin-40* interallelic complementation tests showed *lin-40* to be a complex locus; therefore, an interallelic complementation map for the alleles of *lin-40* was developed. An attempt was made to locate *lin-40* on the physical map developed by Coulson *et al.* (1986; 1988).

MATERIALS AND METHODS

General

A 700 base pair stretch of Tc1 coding region in the plasmid pCeS100 was kindly provided by K. Beckenbach.

Brood sizes and two-factor mapping to unc-46

All *lin-40* alleles used except *e2173* are linked to *unc-46(e177)*. To obtain both two-factor map distances and brood sizes, heterozygous (over wild-type chromosomes) late larval hermaphrodites with appropriate mutations were set up individually and brooded every 12 hours until they were exhausted of progeny. Rose and Baillie (1979) showed that recombination is temperature dependent, therefore all brooding was done at 20C. All late larval and mature F_1 progeny were scored.

Interallelic complementation tests

All *lin-40* alleles were balanced by *eT1(III;V)* and pseudolinked to *dpy-18(III)*. (For a description of the *eT1* balancing system see Rosenbluth and Baillie 1981). The presence of fertile Dpy Unc (Dpy-18 where *e2173* was involved) F_1 progeny, resulting from the cross of strains containing two alleles of *lin-40*, indicated complementation. The absence of the Dpy Uncs indicated failure to complement. Homozygous *eT1* nematodes have the uncoordinated Unc-36 phenotype. The absence of Dpy Unc was accepted only after the plate was scored for all males and a minimum of ten *eT1* males were identified.

Developmental blocking stage

The blocking stages for *dpy-18(III); (let-336 or lin-40) unc-46(V)* were noted at 20C. Some aneuploid progeny of *+/eT1(III);+/eT1(V)* block as eggs and others block as L1 larvae. Therefore, to observe the blocking stages of the early blocking alleles of *let-336* and *lin-40*, *dpy-18/eT1;(let-336 or lin-40) unc-46/eT1* hermaphrodites were mated with *dpy-18/eT1; unc-46/ eT1* males and (1-3) F_1 Dpy Unc hermaphrodites were brooded for six

hours. After 3 days all maturing F_2 were counted and removed. On the fourth day the homozygous *dpy-18; (let-336 or lin-40) unc-46* were counted and on the fourth and seventh day they were observed and their lengths were measured. To ascertain the developmental blocking stages of the late blocking mutations, 20 *dpy-18/eT1(III); lin-40 unc-46/eT1(V)* hermaphrodites were brooded for 6 hours. Three days later, several F_1 Dpy Uncs were transferred to fresh plates for observation and measurment on the fourth and seventh days.

Preparation of genomic DNA

The Tc1 induced recessive lethal alleles of *lin-40* (Clark *et al.* 1990) were balanced in this genetic background: *dpy-18/eT1(III);lin-40 unc-46/let-x(s2165)eT1(V). eT1(III;V)* is a reciprocal translocation that balances *lin-40. let-x(s2165)eT1(V)* (unpublished results) is a recessive early blocking lethal mutation on the balanced region of *eT1(V)*, included to minimize the contribution of wild-type *lin-40* to the DNA preparation. Genomic DNA was prepared by using the method described by Emmons *et al.* (1979) with these two exceptions 1) the worms were grown on high peptone (eight times higher than used in simple nematode growth medium) agarose plates and 2) the worms were not frozen nor ground with a mortar and pestle. Growing worms on agarose plates obviates the need to purify the DNA on a CsCl gradient.

Restriction enzyme digestion

Genomic and plasmid DNAs were digested with restriction enzymes obtained from either Bethesda Research Laboratories (BRL) or from Pharmacia. The buffers used were those recommended by BRL. Restriction digests were always carried out with at least 2 units/microgram of DNA for a minimum of one hour at 38C. All digests were stopped by heating at 68C for 10 minutes.

Agarose gel electrophoresis

Approximately 3 micrograms per strain of *Eco*RI digested genomic DNA was loaded onto a 0.5% agarose gel. Electrophoresis was performed at 25 Volts overnight in 1X TBE buffer [10X=89mM Tris, 89mM boric acid, 2.5mM EDTA (pH8.3)] Maniatis *et al.* (1982).

Southern Transfer

Following electrophoresis, the gel was soaked in four volumes of 0.25M HCl for 30 minutes, rinsed in double distilled water and then soaked in four volumes of 0.5M NaOH, 1.5M NaCl for 30 minutes. Finally, the gel was soaked in transfer buffer (1M ammonium acetate, 0.02M NaOH) for one hour. The DNA was then transferred to nitrocellulose (Schleider and Schuell) by the one directional method described in Davis *et al.* (1980).

Labelling DNA probes

DNA probes were labelled with ^{32}P-dATP either by the nick translation technique of Rigby *et al.* (1977) or by the oligolabelling technique of Feinberg and Vogelstein (1983). The labelling reaction was stopped (10X stop solution: 200mM EDTA, 5%SDS) and put through a G-25 Sephadex (Pharmacia) spin column. The specific activity of the probes was measured and the required amount was filtered with hybridization buffer through a 0.45 micron nitrocellulose filter (Schleider and Schuell).

Isolation of DNA fragments from agarose gels

Digested DNA was run in a 0.5% agarose gel and the desired fragment was cut from the gel. DNA was extracted from the gel slice by using an Elutrap (Schleider and Schuell) apparatus and the protocol provided by the manufacturer.

Construction of lambdaZAP mini libraries

LambdaZAP (Stratagene) contains single stranded bacteriophage sequence that allows excession of a phagemid from the isolated lambda phage. The phagemid can be maintained in *E. coli* as a single or double stranded clone. Approximately 200 nanograms

of isolated *Eco*RI digested genomic DNA combined with one microgram of lambdaZAP DNA was ligated and packaged according to the Stratagene protocol.

Hybridization of probes to DNA filters

Prehybridization was done for at least one hour in hybridization solution (5 X SSPE, 0.3% SDS and 2.5 X Denhardt's) at 68C. Hybridization of the filters with DNA probe was done overnight (at least 16 hours) at 68C. Washing was done with 1 X SSPE at 68C.

Screening Genomic libraries

To isolate individual plaques three successive screens were done. The phage titres were approximately: 10^5/filter for the first screen, 500 for the second and 20-50 for the third. Appropriate oligolabelled DNA probes were used to hybridize to filters of the libraries.

RESULTS

lin-40(s1352) is located 4.7 m.u. (3.4-6.2) (95% confidence limit) to the left of *unc-46* on LGV(left) (Figure 17). Sixteen recessive (lethal to *mel*) alleles of *lin-40* have been analyzed. Interallelic complementation tests were performed and the results indicate that *lin-40* is a complex locus with at least five complementation classes. The five classes have been named A, B, C, D and E. A contains 1 allele (*s1506*), B contains three alleles (*s1593*, *s1669* and *s1675*), C contains four alleles (*s1053*, *s1611*, *s1634* and *s1704*), D contains one allele (*e2173*) and E contains seven alleles (*s1345*, *s1351*, *s1352*, *s1358*, *s1360*, *s1373* and *s1916*). Class A complements classes B, C and D but fails to complement E. B complements C but fails to complement D and E. C fails to complement D and E. D fails to complement E. A complementation map based on these results is shown in Figure 18.

The developmental blocking stages due to the various homozygous *lin-40* alleles range from early-mid larval lethal to maternal effect larval lethal. The class A and E alleles (*s1506*, *s1345*, *s1351*, *s1352*, *s1358*, *s1360*, *s1373* and *s1916*) all block development at early-mid larval stages. The blocking stages of the class B, C and D alleles (*s1593*, *s1675*, *s1669*, *s1611*, *s1634*, *s1053*, *s1704* and *e2173*) range from late larval lethal or sterile adult to maternal effect larval lethal. All of the class B, C and D alleles cause a protruding vulva. Any combination of *lin-40* alleles that fail to complement, or any allele over a deficiency, appear to cause blocking at the same stage as that of the homozygous later blocking allele.

Nematodes homozygous for some of the *lin-40* alleles were observed through a microscope fitted with Nomarski optics. Homozygous *s1053* hermaphrodites are slow developing sterile adults. The gonads form (but there is some gonadal wandering) and fill with a fluid (possibly yolk fluid). No germ-line forms and meiosis does not occur. Some

uterus cells appear abnormal and are possibly distal tip cells. There appear to be the normal number of cells in the vulva but it is possible that the attachment of the vulva to the uterus is abnormal. The vulva protrudes and there is no vulva opening. In the homozygous *s1053* male, the spicules are abnormal but all appear to be present. The tail is deformed and the gonad does not migrate. There is no yolk fluid in the male.

Homozygous *s1669* hermaphrodites have a relatively normal looking gonad but the vulva protrudes and the vulva appeared to be plugged. The eggs form and contain a fairly advanced larval worm but the eggs do not hatch. The cell division in the eggs appear to be abnormal.

Homozygous *s1675* hermaphrodites have a relatively normal gonad but the vulva is plugged and protrudes. The cell division in the eggs may be slow but a few eggs mature and hatch inside the parent. Larval progeny have not been observed externally.

Homozygous *e2173* hermaphrodites have short incompletely formed gonads that do not reflex. All of the somatic cells are probably present. There appear to be 1-2 germ-line nuclei/arm. The vulva protrudes and does not have an opening.

Development of homozygous *s1506* hermaphrodites is blocked at an early-mid larval stage; the larval worms look dumpyish and have an abnormal tail. No gonadal development was observed.

In Table 9 the results of the scoring for brood sizes are listed. N2 (wild-type) were scored by Hodgkin (1988). As some of the mutations listed in Table 9 are early blocking recessive lethals, only 3/4 of their progeny would have been scored. For N2, *dpy-18/+;unc-46/+* and the strains containing *lin-40* alleles that are blocked late in development *(e2173, s1675, s1669, s1593, s1053, s1611 and s1634)*, all the progeny were scored and 3/4 of the scored values are listed in Table 9. The strain containing *lin-40(s1704)*, which is a late blocking allele, also has a linked early blocking recessive lethal mutation. Therefore the numbers listed for *s1704* are the totals scored. It was

noticed during scoring that the strain containing *lin-40(s1506)* had an early blocking lethal mutation on LGIII (there were very few F_1 Dpy-18s), therefore 4/3 of the number scored are listed in Table 9.

sDfs 27,42 and *50* delete *lin-40* while *sDfs 34,40* and *70* break close to *lin-40* but do not delete it (*sDf40* also deletes *let-336*) (Figure 17). Hermaphrodites from strains containing the three deficiencies listed (Table 9) that delete *lin-40* average 40% larger brood sizes than hermaphrodites from the two strains containing the deficiencies (*sDfs 34* and *70*) that do not delete *lin-40*. T-tests (95% confidence) show that the differences are significant. T-tests also show that the mean self-progeny brood sizes for heterozygous *sDf42* and *sDf50* animals are significantly larger than the mean self-progeny brood sizes for *dpy-18/+;unc-46/+* individuals. The mean self-progeny brood sizes of hermaphrodites of strains containing heterozygous *sDfs 34* or *70* are significantly smaller than the self-progeny brood sizes of *dpy-18/+;unc-46/+*. The small brood sizes is attributible to dominant effects due to the deficiencies. To test this, two to three *dpy-18/+;sDf unc-46/+* animals were brooded for six hours and the eggs laid were counted. Twenty-four hours later the larvae and unhatched eggs were counted, and two days later the maturing Dpy Uncs were counted and removed. For *sDf34*: 49 eggs were laid, 8 did not hatch, 34 matured to adulthood, and 7 stopped development as early larvae. *sDf34* deletes *emb-29* (Figure 17) and homozygous *emb-29* die as eggs; therefore, homozygous *sDf34* can not mature past the egg stage. The arrested larvae must have been heterozygous for *sDf34*, and they account for the low self-progeny brood sizes of *sDf34* heterozygotes. Similar results were obtained for *sDf70*.

Four of the strains containing *lin-40* alleles (*e2173*, *s1351*, *s1506* and *s1675*) have mean self-progeny brood sizes significantly larger (T-test) than *dpy-18/+;unc-46/+* and, although few broods were scored, several others had large broods much bigger than the largest in the range of the 19 N2 and *dpy-18/+;unc-46/+* hermaphrodites scored (Table 9).

A second gene, *let-336*, which maps 14.8 (12.4-17.2) m.u. to the left of *unc-46* (Figure 17), also shows the phenomenon of dominantly causing an increase in the number of self-progeny from hermaphrodites. Hermaphrodites heterozygous for either *sDf39* (which deletes *let-336*) or *let-336(s1413)*, show mean self-progeny brood sizes significantly (T-test) larger than hermaphrodites containing just *dpy-18/+;unc-46/+*.

Genomic DNA was prepared from the following strains: wild-type N2 (K. Beckenbach pers. comm.); strains containing the mobilization of Tc1 induced *lin-40* alleles *s1345*, *s1351*, *s1352*, *s1358*, *s1360* and *s1373* (all in the genetic background described in Materials and Methods); and BC2219 which is heterozygous for *dpy-18/eT1(III);unc-46/let-x(s2165)eT1(V)*. DNA from each of the above strains was digested with *Eco*RI, electrophoresed and blotted by Southern transfer. Nick translated ^{32}P-dATP labelled Tc1 DNA was used to probe the digested DNA. Three of the Tc1 induced *lin-40* alleles (*s1351*, *s1358* and *s1373*) showed a common new band of approximately 7.5 kb. A 6.5 - 8.0 kb lambdaZAP library containing the new Tc1 band from *s1351* was constructed and a phage containing the new Tc1 band was isolated.

The plasmid containing the new Tc1 band and its flanking DNA was digested with *Eco*RV to remove the Tc1 sequence (it also removes an *Eco*RI site in the polylinker). The resulting plasmid, when cut with *Eco*RI and *Hind*III, yielded 0.9 and 1.2Kb fragments. The 1.2Kb fragment was used to probe *Eco*RI digested wild-type N2 DNA and the *Eco*RI digested genomic DNA from each of the Tc1 induced heterozygous *lin-40* strains (*s1345*, *s1351*, *s1352*, *s1358*, *s1360* and *s1373*). Two bands approximately 1.6Kb apart were in the *s1351*, *s1358* and *s1373* containing DNAs while only one band was found in the N2, *s1345*, *s1352* and *s1360* DNAs (see Figure 19). The N2 band was the same length as the shorter band from the DNA containing *s1351*, *s1358* and *s1373*. The complete Tc1 DNA is about 1.6Kb long; therefore, the isolated DNA flanking the new Tc1 is single copy and

s1351, *s1358* and *s1373* have a common Tc1 insertion. It is important to note that all the Tc1 induced *lin-40* alleles originated independently (Clark *et al.* 1990).

The 1.2 Kb flanking DNA was used to probe T. Snutch's (1984) Charon 4 N2 library #1. This library was constructed using N2 genomic DNA that was partially digested with *Eco*RI and size selected for 15-23 Kb DNA fragments. A Charon 4 phage (BC#S1001) was identified and isolated and the inserted DNA, putatively containing *lin-40*, sent to A. Coulson and J. Sulston (Cambridge, England) to be located on the physical map of *C. elegans*. The position of BC#S1001 is shown in Figure 20.

Figure 17

map of the *let-336-lin-40* region of LGV(left)

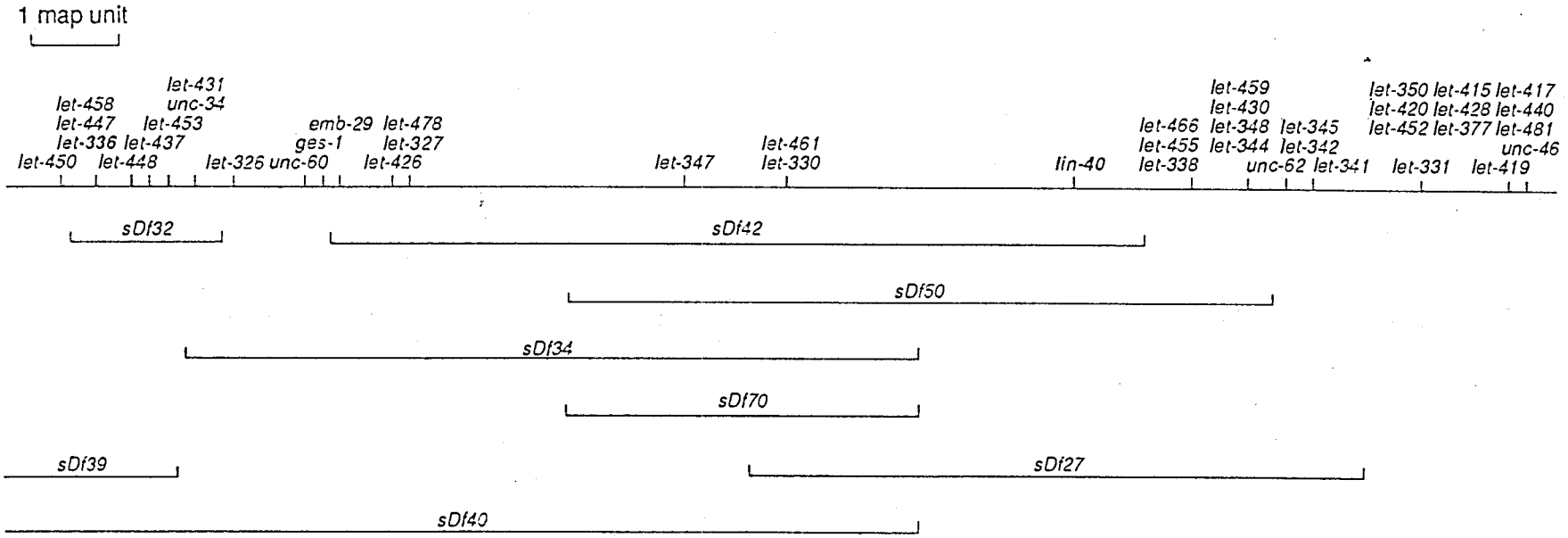

1 map unit
let-431
unc-34
let-458
let-447
let-453
let-336
let-437
emb-29
ges-1
let-478
let-327
let-450
let-448
let-326
unc-60
let-426
let-347
let-461
let-330
lin-40
let-459
let-430
let-466
let-348
let-345
let-455
let-344
let-342
let-338
unc-62
let-341
let-350
let-415
let-417
let-420
let-428
let-440
let-452
let-377
let-481
unc-46
let-331
let-419
sDf32
sDf42
sDf50
sDf34
sDf70
sDf39
sDf27
sDf40

Figure 18

lin-40 complementation map based on the results of interallelic complementation testing of the 16 *lin-40* alleles.

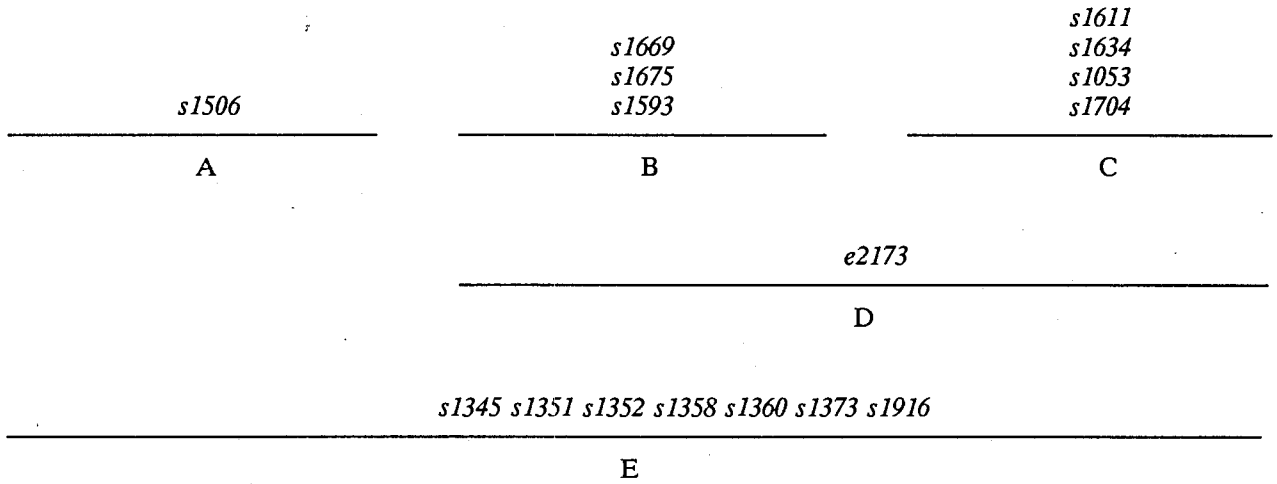

s1506
A
s1669
s1675
s1593
B
s1611
s1634
s1053
s1704
C
e2173
D
s1345 s1351 s1352 s1358 s1360 s1373 s1916
E

TABLE 9

(Self-progeny brood sizes of heterozygous hermaphrodites)

MUTATION	NUMBER OF BROODS COUNTED	AVERAGE BROOD SIZE	RANGE	CLASS
N2(wild-type)[a]	12	246	205-280	
dpy-18/+;unc-46/+	7	249	216-275	
sDf27	9	279	192-337	
sDf42	8	272	240-314	
sDf50	4	283	262-294	
sDf40	6	216	196-236	
sDf34	4	199	174-238	
sDf70[b]	7	199	164-229	
lin-40(s1506)[c]	5	273	252-309	A
lin-40(s1675)[d]	11	275	215-322	B
lin-40(s1669)[d]	6	256	231-294	
lin-40(s1593)[d]	5	234	204-262	
lin-40(s1704)[e]	17	239	207-255	C
lin-40(s1053)[d]	12	235	199-258	
lin-40(s1611)[d]	5	233	212-277	
lin-40(s1634)[d]	2	217	200-233	
lin-40(e2173)[f]	6	299	279-337	D
lin-40(s1351)	8	278	251-338	E
lin-40(s1360)	3	270	213-314	
lin-40(s1916)	5	262	246-276	

TABLE 9 (cont)
(Self-progeny brood sizes of heterozygous hermaphrodites)

MUTATION	NUMBER OF BROODS COUNTED	AVERAGE BROOD SIZE	RANGE	CLASS
lin-40(s1373)	8	251	209-325	E
lin-40(s1352)	8	238	215-255	
lin-40(s1358)	4	238	229-261	
lin-40(s1345)	6	213	194-234	
sDf39	6	285	256-329	
let-336(s1413)	4	286	248-322	

[a] N2 (wild type) were scored by J. Hodgkin (1988). Most mutations listed in this table are early blocking recessive lethals and therefore 1/4 of the progeny of hermaphrodites heterozygous for the recessive lethal block early and were not scored, but N2 is wild type and all the progeny were scored. To compare the mean brood size of N2 to the mean brood sizes of the hermaphrodites heterozygous for the other mutations, 3/4 of the total N2 progeny counted are listed in the table.

[b] *sDf70* scored by H.I. Stewart (pers. comm.).

[c] *s1506* blocks early, while scoring for brood size it was noticed that the strain used also had an early blocking lethal mutation on LGIII, therefore the numbers listed are 4/3 of the numbers scored.

[d] These alleles of *lin-40* are all late blockers, therefore the numbers listed are 3/4 of the numbers scored.

[e] The strain used that contains *s1704* also contains a recessive early blocking lethal on the same chromosome. Therefore even though *s1704* causes late blocking of development all of the progeny scored are listed in the table.

[f] *lin-40(e2173)* was scored by R.E. Rosenbluth (pers. comm.). *e2173* causes a recessive sterile adult phenotype and all the progeny of heterozygous *e2173* hermaphrodites were scored, therefore for comparison, 3/4 of the numbers scored are listed in this table.

Figure 19

N2 is wild-type, *s1351*, *s1358*, *s1373*, *s1352*, *s1345* and *s1360* are recessive lethal alleles of *lin-40* induced by the mobilization of the transposon Tc1 (Clark *et al.* 1990). A new approximately 7.5 Kb band was noted in *Eco*RI digested genomic DNA from strains heterozygous for *s1351*, *s1358* and *s1373* when probed with Tc1 DNA. The extra *Eco*RI band was isolated from the strain containing *s1351*. The Tc1 was removed and flanking DNA was used to probe N2 and heterozygous *s1351*, *s1358*, *s1373*, *s1352*, *s1345* and *s1360* bearing strains. The autoradiograph shows that the new band 7.5Kb (upper band) is common to *s1351*, *s1358* and *s1373*. Because the six strains arose independently (Clark *et al.* 1990) and three strains show a common Tc1 band, *lin-40* has probably been cloned.

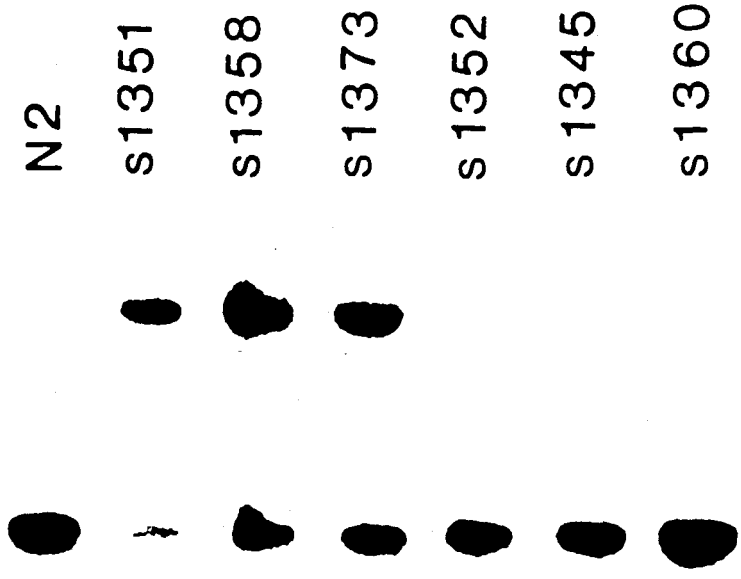
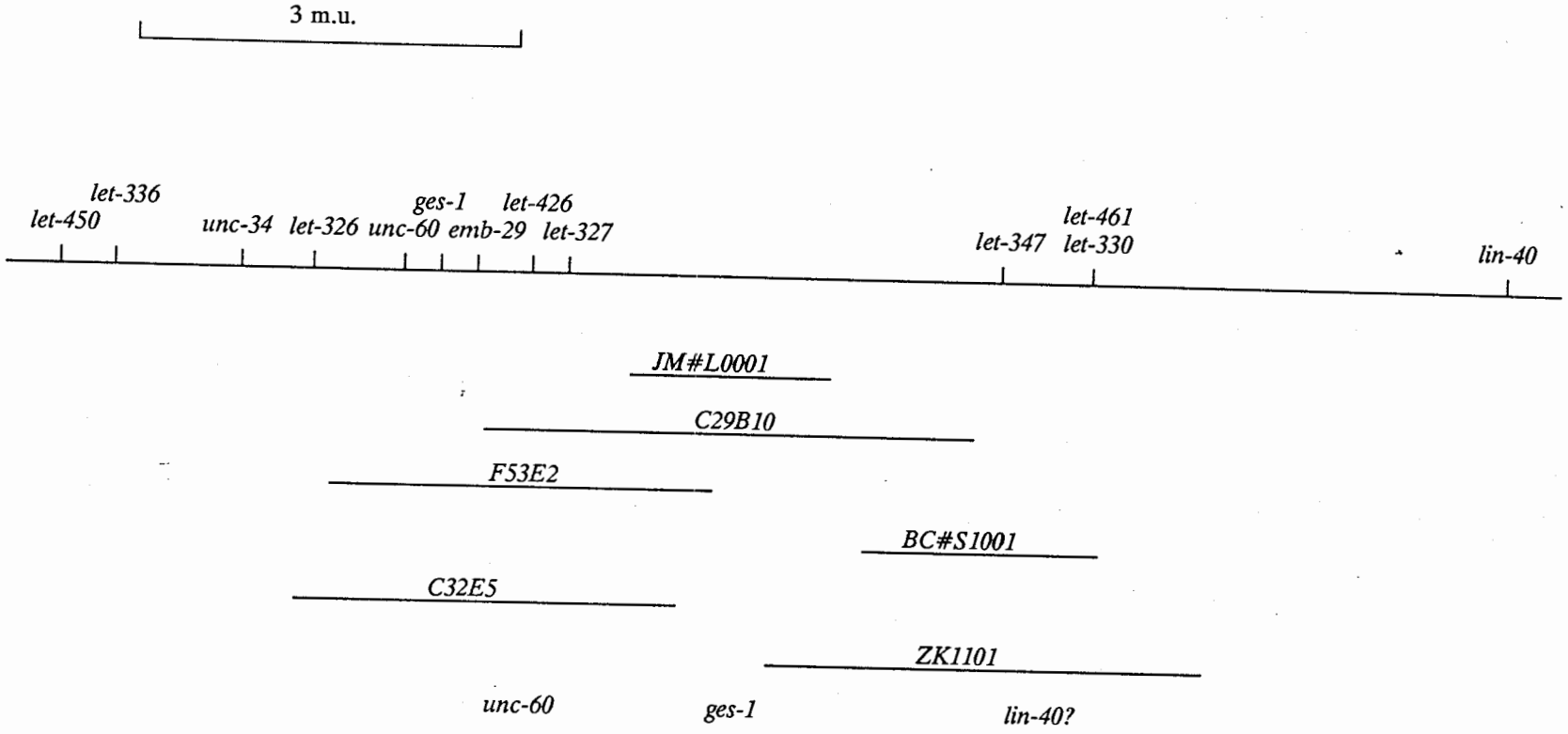

N2
s1351
s1358
s1373
s1352
s1345
s1360

Figure 20

Physical map (A. Coulson and J. Sulston pers. comm.) and genetic map of LGV(left). *ges-1* is in JM#L0001 (J. McGhee pers. comm.). M. Wakarchuk located *unc-60* in cosmid F53E2. BC#S1001 contains DNA that hybridizes strongly to the new *Eco*RI band that was isolated from the strain containing *lin-40(s1351)*.


3 m.u.
let-336
let-450
unc-34
let-326
ges-1
unc-60
emb-29
let-426
let-327
let-461
let-347
let-330
lin-40
JM#L0001
C29B10
F53E2
BC#S1001
C32E5
ZK1101
unc-60
ges-1
lin-40?

DISCUSSION

XX hermaphrodites produce approximately 320 sperm before switching to exclusive oocyte production. As the number of self-progeny from hermaphrodites is limited by the number of sperm produced, mutations that result in the increase in the number of self-progeny probably identify genes that play a roll in the switch from spermatogenesis to oogenesis. In this Thesis two essential genes, *let-336* and *lin-40*, that have mutations that dominantly cause an increase in the mean number of self-progeny have been identified. One weak allele of *tra-2* (Hodgkin 1988) and gain-of-function alleles of *fem-3* (Barton *et al.* 1987) cause similar effects.

The switch to oogenesis requires that *fems 1,2,3* and *fog-1* cease to direct spermatogenesis. *tra-2* activity down-regulates the *fem* genes and *fog-1*, and also has control over the genes responsible for somatic sex differentiation. The phenotype of null *tra-2* alleles is the complete masculinization of the germ-line and incomplete masculinization of the soma (Hodgkin and Brenner 1977). Loss-of-function *fem-3* alleles cause the feminization of soma and germ-line in both XX and XO animals (Hodgkin 1986), but temperature sensitive gain-of-function *fem-3* alleles that only affect the germ-line of XX animals have been isolated and analyzed by Barton *et al.* (1987). Temperature shift experiments showed that germ-line sexual commitment is a continuing process; spermatogenesis can resume after oogenesis has started and the onset of oogenesis can be delayed.

The brood sizes in this study were determined by counting the mature self-progeny produced by hermaphrodites. Counting mature progeny is subject to biases that lead to low scored brood sizes. Any degree of dominant lethality due to the mutation would lower the scored brood size. The mutations that were being analyzed for their effect on brood size arose in screens where the whole genome was subject to mutations and any slightly

deleterious mutations elsewhere in the genome could also result in a reduction of the number of mature progeny. These biases could be partially eliminated by scoring the eggs laid. Scoring eggs was not done because the procedure is labourious. The P_0 hermaphrodites could have one of two genotypes: *dpy-18/+;(let-336 or lin-40) unc-46/+* or *+/eT1(III);+/eT1(V)* and only by observing maturing F_1s could the broods of *+/eT1(III);+/eT1(V)* hermphrodites be eliminated.

One of the three *let-336* alleles (*s1413*) was used to determine the dominant effect of the essential gene *let-336* on the self-progeny brood sizes of hermaphrodites. The mean brood size was about 15% larger than expected for wild-type (including *dpy-18/+;unc-46/+* which is not significantly different from the wild-type) and was significantly different from wild-type brood sizes by a T-test (95% confidence limit). The brood size due to the deficiency *sDf39* was also significantly larger than expected (15%). *sDf39* deletes *let-336* and at least six other essential genes (Figure 17). The average brood size due to *sDf40* (which deletes *let-336*) was not exceptionally large, although it was larger than the sizes due to either *sDf34* or *sDf70* (which do not delete *let-336*). *sDf40* is larger than either *sDf34* or *sDf70* and deletes all the known genes that either of them do. The absence of an exceptionally large brood size is probably attributable to the large size of *sDf40*. There may be some dominant effects due to having that large region hemizygous. Three deficiencies that delete *lin-40* (*sDfs 27, 42* and *50*) were analyzed for their effect on self-progeny brood size. All three caused increased broods sizes similar to the brood sizes caused by *sDf39*. The fact that deficiencies that delete *let-336* or *lin-40* cause increased brood sizes implies that the switch from spermatogenesis to oogenesis is affected by dosage and that neither *let-336* nor *lin-40* are regulated to compensate for the deleted genes.

The sixteen alleles of *lin-40* were tested for their effect on brood size. Four alleles (*s1506*, *s1675*, *e2173* and *s1351*) caused significantly larger brood sizes than expected (Table 9), and although few broods were scored, other alleles (*s1669*, *s1360* and *s1373*)

had some broods larger than any of the 19 wild-type broods. The alleles *s1593*, *s1611*, *s1634* and *s1345* were in unhealthy strains. For the region around *lin-40* two types of deficiencies are listed in Table 9, those that delete *lin-40* and those that do not. Because of the similarity of brood sizes due to deficiencies of each type, the comparison of brood sizes due to the deficiencies should more accurately reflect the effect of haploid *lin-40* on self-progeny brood size than the comparison of wild-type brood sizes to brood sizes caused by the various *lin-40* alleles. Hermaphrodites containing deficiencies that delete *lin-40* average 40% larger broods than hermaphrodites containing deficiencies that do not delete *lin-40*. The small brood sizes caused by deficiencies that do not delete *lin-40* are not due to the deletion of genes not deleted by the other deficiencies. *sDf70* is completely contained in *sDf42* (Figure 17). *sDf42* deletes *lin-40* but *sDf70* does not. The mean brood size of heterozygous *sDf42* hermaphrodites is 40% larger than the mean brood size of heterozygous *sDf70* hermaphrodites.

lin-40 has a complex interallelic complementation pattern. Sixteen recessive alleles have been identified and placed into classes according to the results of interallelic complementation tests (Figure 18). It is possible that *lin-40* represents three genes: the class E alleles are all deficiencies that delete all three genes and the class D allele is a deficiency that deletes two of the genes. This is unlikely because of the similarity of phenotypes between classes and the fact that alleles in different classes have similar effects on the self-progeny brood sizes of hermaphrodites and that none of the ends of the class E alleles are "fuzzy", *i.e.* they do not fail to complement any known genes to the left or right of *lin-40*. Although other genes in *C. elegans* show interallelic complementation [examples include *unc-15* (Rose and Baillie 1981); *cha-1-unc-17* (Rand 1989); *bli-4* (K. Peters pers. comm.)] none have as complex a pattern as *lin-40*. The complex interallelic complementation pattern of *lin-40* is reminiscent of the *maroon-like* locus in *Drosophila melanogaster* (Chovnick *et al.* 1969; Finnerty *et al.* 1970; Finnerty and Chovnick 1970),

where five complementation classes were also identified. The enzyme level analysis and fine structure analysis of *maroon-like* led to the conclusion that it is a single gene.

The developmental blocking stages due to the various *lin-40* alleles (all are recessive) range from early-mid larval to maternal effect larval lethal. None of the hermaphrodites homozygous for any *lin-40* allele has been rescued by mating. The blocking stages appear to be consistent with the interallelic complementation pattern. Class A and E alleles are all early-mid larval blockers in both hermaphrodites and males. Class B and C have only late larval or sterile adult and maternal effect alleles. The one class D allele causes late larval lethality or sterile adults. All of the B, C and D alleles cause a protruding vulva. Some of the hermaphrodites have poorly developed gonads while others have relatively normal looking gonads. The homozygous males' tails and gonads are abnormal for some of the alleles. Animals heterozygous for any two of the various alleles (when they fail to complement) invariably show the phenotype of the later blocking allele and it does not matter which allele came from which parent. Class A, B, D and E alleles dominantly cause an increased self-progeny brood size but none of the four class C alleles do. The *lin-40* products affect both sexes and the different classes of alleles have different effects. The effects range from early-mid larval lethality to gonadal and secondary sexual abnormalities, to dominant germ-line switch delays, to maternal effect lethality. This implies that the *lin-40* product is not only needed in both sexes but acts at several developmental stages and has effects in several different tissues.

The interallelic complementation pattern implies a physical ordering of the classes of *lin-40* alleles. The deduced order is reflected in Figure 18. The fact that the class D allele complements class A but fails to complement classes B and C suggest that D is a polar mutation transcriptionally down-stream from class A. The fact that class A alleles are rare (1 allele) and class E alleles are common (7 alleles) supports this ordering. Null alleles that stop function downstream (class E) should be more common than alleles that

negate one exon but do not effect exons down-stream from it (class A). The relative ordering of classes B and C is more ambiguous. The reason they were placed in the order shown in Figure 18 is that alleles of the classes A, B and D cause increased brood sizes but none of the class C alleles do. The classes that have similar effects were placed in proximity.

It is interesting that an allele of *let-336* and some of the alleles of *lin-40* that dominantly cause an increase in the number of self-progeny from hermaphrodites also cause early larval lethality when homozygous. This suggests that there is a link between early larval development and the much later switch from spermatogenesis to oogenesis. The early larval lethal phenotype is probably not due to antimorphic alleles because, generally, antimorphs are rare and all three alleles of *let-336* and several alleles of *lin-40* are early larval blockers. The switch from spermatogenesis to oogenesis can not be the result of antimorphic or neomorphic alleles because wild-type over deficiencies for both *let-336* and *lin-40* cause increased brood sizes.

The 16 alleles of *lin-40* were induced by various mutagens, but six of the alleles were induced by Tc1 mobilization. The molecular tagging of specific genes using Tc1 has proven useful (see Greenwald 1985; Moerman *et al.* 1986 for examples; for a review on Tc1 see Moerman and Waterston 1989). *mut-4(I)* has been shown to induce Tc1 mobilization (Mori *et al.* 1988b). Clark *et al.* (1990) used *mut-4* to induce recessive lethal mutations over the 49 m.u. of LGIV(right) and LGV(left) that are balanced by the translocation *nT1* (Ferguson and Horvitz 1985; Clark *et al.* 1988). 24 out of the 28 recessive lethals recovered mapped to LGV(left), and 6 out of the 15 LGV(left) lethals analyzed are putative point mutations in *lin-40* and several of the others are putative tip deficiencies that broke fairly close to *lin-40*. Thus, *lin-40* and the region around *lin-40* are major targets for mutagenesis induced by the mobilization of Tc1. This clearly showed that transposon induced mutagenesis can have strong biases.

Genomic DNA was prepared from the six strains containing Tc1 induced alleles of *lin-40*. The DNA was *Eco*RI digested, electrophoresed and probed with Tc1 DNA. A new Tc1 containing *Eco*RI band was identified that is common to the three strains containing *s1351*, *s1358* or *s1373*). DNA flanking the Tc1 in the new band was isolated and used to probe genomic DNA from the strains containing Tc1 induced alleles of *lin-40*. The flanking DNA hybridized to the common new bands (see Figure 19). The fact that all the Tc1 induced alleles of *lin-40* arose independently (Clark *et al.* 1990) and that the new *Eco*RI band is common to three strains each containing a different *lin-40* allele, strongly suggest that the new band coincides with *lin-40*. A Charon 4 phage DNA (BC#S

1001) containing N2 genomic DNA corresponding to the new band was sent to Cambridge for positioning on the physical map (A. Coulson and J. Sulston pers. comm.). The position of BC#S1001 is shown in Figure 20.

If BC#S1001 does contain *lin-40* then the physical position of *lin-40* (Figure 20) is surprisingly far to the left (close to *ges-1* and *unc-60*, even though it is as much as 8-9 m.u. from *unc-60*). Although the genetic distance between *unc-60* and *lin-40* has not been measured directly, the position of *lin-40* on the map (Figure 17) is consistent with mapping results of McKim *et al.* (1988a); Rosenbluth *et al.* (1988); Rosenbluth *et al.* (1990), where various genes to *lin-40*'s left had been two-factor and three-factor mapped among themselves and against genes to the right of *lin-40*. Results from identifying the number of coding elements in other regions of known length both physically and recombinationally, suggest that gene spacing along the chromosome and recombinational distances do not have a strong correlation (Greenwald *et al.* 1987; Prasad and Baillie 1989; Starr *et al.* 1989). Greenwald *et al.* (1987) showed that the *eP6-lin-12(III)* interval has 930 kb/m.u. and the *lin-12-eP7(III)* region has 830 kb/m.u.. This compares to the average value of 333kb/m.u. for *C. elegans* [the calculation was based on dividing the genome size of $1X10^5$ kb (J.E. Sulston pers. comm.) by the number of map units (300) (Edgley and Riddle

1990)]. Prasad and Baillie (1989) identified 9 coding regions in 150 kb that spans approximately 0.2 m.u. near *unc-22(IV)*. This yielded 750 kb/m.u.. Starr *et al.* (1989) used DNA polymorphisms across the LGI gene cluster to position more than 1000kb of DNA with respect to the genetic map. They noted that the amount of DNA/m.u. was not constant across the region and peaked at about 950kb/m.u.. Assuming that the left end of the *unc-60* bearing cosmid (F53E2) (M. Wakarchuk pers. comm.) and the right end of BCS1001 are a maximum of 100 kb apart, my data suggests that there may be as few as 12 kb/m.u. in the *lin-40* region, that is, the region has a relatively high rate pf recombination. The region around *lin-40* has few identified genes; this again may indicate relatively little DNA per m.u. and that *unc-60* is physically close to *lin-40*.

A few attempts to rescue *lin-40* by the microinjection of appropriate DNA have been tried, but no rescue has resulted. To determine whether tyhe inserted Tc1s are modified by mut-4, a total of 10,000 nematodes containing *mut-4* and Tc1 induced alleles of *lin-40* have been screened but no reversion has been detected.

lin-40 resides in a high recombination/gene poor region of LGV(left). The region is highly susceptible to mutagenesis by the mobilization of the transposon Tc1 (Clark *et al.* 1990). Some of *lin-40*'s neighbouring genes are very interesting. To the left of *lin-40* are *let-330* and *let-461*. *let-330* is by far the largest target for EMS mutagenesis on LGV(left) (Section 4). *let-461* has recessive early larval lethal alleles but it also has alleles that are semi-dominant suppressors of *lin-12* dominant alleles (J. Thomas pers. comm.), and alleles that have the same phenotype as *lin-12 glp-1* double mutants (J. Kimble and E. Lambie pers. comm.). To the immediate right of *lin-40* are *lets 338, 455* and *466*. Both alleles of *let-338* are recessive mid-larval lethals (Rosenbluth *et al.* 1988); both alleles of *let-455* are recessive for early larval lethality (Section 4), and both alleles of *let-466* are recessive and cause adult sterility (Section 4).

It is possible that there are several genes that affect brood size. The effect on brood size of *lin-40* and *let-336* was noticed when the two factor mapping of mutations on the *eT1* balanced region of *LGV(left)* was done. This region comprises approximately 7% of the recombination distance in the genome. Large brood sizes due to deficiencies on LGII have also been noted by A. Brooks (pers. comm.) in T. Johnson's laboratory. These results suggest that the switch from spermatogenesis to oogenesis is affected by or under the control of many genes.

GENERAL DISCUSSION

In this thesis I have focused on the genetic analysis of the left half of linkage group V in *Caenorhabditis elegans*. By the mapping of deficiency breakpoints, I have increased the number of zones that subdivide LGV(left). These zones facilitated the mapping of mutations to location on LGV(left). I have mapped a large set of putative point mutations to zones and to genes. I have also done the initial characterization of the function(s) of some of the essential genes that are on LGV(left).

LGV(left) is the largest region in any higher eukaryote for which an attempt to saturate for essential genes has been made. Rosenbluth *et al.* (1988) had used chromosomal rearrangements to subdivide LGV(left) into 22 zones, I have increased that number to 36 by mapping formaldehyde and Tc1 induced deficiencies of the region.

While analyzing the set of formaldehyde induced mutations, I developed a dose-response curve and recommended a dose of 0.1% for formaldehyde mutagenesis. I also described the types of mutations induced by formaldehyde. The mutations induced were generally chromosomal rearrangements and the deficiencies induced averaged about one half the size of those induced with gamma irradiation. Thus formaldehyde is useful for inducing relatively small deficiencies.

The Tc1 mutations were induced to generate a set of tagged genes that will be useful in the correlation of the genetic and physical maps of the region. Three genes have been putatively tagged with Tc1 (*let-448*, *let-449* and *lin-40*). Tc1 mutagenesis gave two surprising results: 1) *lin-40* was extremely susceptible to Tc1 mutagenesis. At least three of the six *lin-40* Tc1 alleles were due to Tc1 insertions. 2) The region between *let-330* and *lin-40* "breaks" frequently, when Tc1 is mobilized, resulting in putative tip deficiencies. These data could indicate that Tc1 uses a homolog dependent repair mechanism similar to the one Engels *et al.* (1990) proposed for P element repair in *D. melanogaster*. The Tc1

mutations were generated over a translocation balancer (*nT1*), and since balancer translocations are thought not to pair in *C. elegans* (Rosenbluth and Baillie 1981), the homolog would not be available to effect repair of resident Tc1 excisions. Therefore Tc1 excisions could result in tip deficiencies.

To understand the mechanics of disjunction and recombination for *C. elegans*' chromosomes, it is necessary to identify the sites necessary for those functions. Chromosomal rearrangments are useful tools for detecting those sites and for developing an understanding their functions. While two-factor mapping deficiencies on LGV(left) with respect to *unc-46*, it was noted that some deficiencies caused suppression of recombination. R.E. Rosenbluth and I did a joint study of this phenomenon. Upon analyzing the effects further we established that the deficiencies fell into two classes of suppressors ("major" and "minor"); all the internal deficiencies studied were major inhibitors. The major inhibitors showed polarity of inhibition, with effectively no inhibition on their left but nearly complete inhibition on their right. From the effects of the deficiency heterozygotes on recombination we developed a model for recombination pairing in *C. elegans*, and postulated a site near the left end of the chromosome (in addition to the pairing initiation site that is at the right end) that was necessary to initiate pairing for recombination. In the model we proposed that there are two types of sites needed for pairing for recombination "initiation sites" and "alignment sites". Intimate pairing for recombination starts at and spreads from each initiation site. Initiation sites are analogous to the pairing points in *D. melanogaster* mapped by Hawley (1980). The site postulated to exist near the left terminus of LGV is an initiation site. Another site must exist in LGV(right) to account for recombination occurring in that region in *eT1(III;V)* translocation heterozygotes. The pairing process spreads from each initiation site by "buttoning-up" the homologs at sequential alignment sites. In the presence of a heterozygous minor inhibitor, there are no homologous left-end initiation sites. Pairing initiates only in LGV(right), proceeds toward

the left, aligns the homologs in a correct manner (*i.e.*, homologous regions remain in register) and no major inhibition of recombination occurs. In the presence of a heterozygous major inhibitor (which is a deficiency does not delete the initiation site), pairing proceeds from both ends of LGV. To the left of the deficiency, homologous regions remain in register and recombine normally. But to its right, the sequential buttoning of alignment sites cause a misalignment. Heterologous regions become aligned, thus inhibiting recombination. This misalignment continues until a region is reached whose alignment was controlled by pairing that initiated at the right end. To explain the disappearance of inhibition in the *ama-2* region (near the center of the chromsome), it is proposed that the distance at which pairing occurs from each end is such that alignment of the *ama-2* region is mainly controlled from the right and therefore remains normal, while the *unc-62* region (on the left) is misaligned because this region is still controlled by pairing that initiated at the left end. There are a minimum of three alignment sites in LGV(left): one in the *sDf32* region, another in the *let-330* region and a third between *sDf27* and *dpy-11*. Stewart *et al.* (1990) have recently identified an internal deficiency (*sDf75*) that is not a major inhibitor of recombination. A crude estimate of the length of chromosome deleted by *sDf75* is 2 m.u.. This observation may be explained by proposing that *sDf75* does not delete an alignment site, suggesting that the alignment sites are at least 2 m.u. apart in this region.

Having characterized the new rearrangements and establishing methods of Tc1 tagging essentials, I undertook an analysis of essential genes on LGV(left). My intention was to identify approximately 70% of the essential genes in the region. In order to do this I screened 10,900 chromosomes treated with 0.012M EMS. From this screen I isolated 194 recessive lethal mutations which mapped to LGV(left). I mapped 166 (my EMS set): one was identified as a deficiency (possibly spontaneous) the rest mapped to specific genes. The new genes identified from my EMS set and the genes I have identified in other screens

total 50. This brought the total number of identified essential genes on LGV(left) to 101. Sixty of the essentials have two or more alleles. Using a double reciprocal plot of the number of genes identified versus the number of mutations analyzed, I estimated that there were 147 essential genes on LGV(left) and 3500 in the entire genome. I have shown that *let-330* is the largest target for EMS mutagenesis (twice as large as the next biggest target) on LGV(left). I identified recessive lethal alleles for three genes which were first identified by alleles that caused visible phenotypes. They include the first lethal allele of *unc-60* and several alleles of *unc-70* and *rol-3*. I showed that a allele of a gene (*let-412*) in the same zone as *rol-3* has a Rol phenotype. From the paradoxical results that I obtained for some alleles of two genes (*let-330* and *let-341*) over deficiencies I deduced that the two genes are tagged by deficiency breakpoints. I attempted to identify the approximate blocking stage for all the alleles identified from my EMS set, I also tried to identify all the temperature sensitive and male rescuable mutations from my EMS set.

If I assume that LGV(left) is representative of the genome and that the genes identified reflect the proportion of genes needed at the different developmental stages then the proportion and minimum number of genes needed at each developmental stage can be calculated. From these assumptions, wild-type genes needed in the zygote for normal egg development consist of 8% (300) of all the essential genes. To get beyond early larval development requires 50% (1750) of the essential genes. To get past the mid-larval stage requires 17% (580) of the essential genes. To get past late larval development requires 2% (73) of the essential genes. 16% (550) of the essential genes at needed to make fertilized eggs. These include the genes needed for the formation and development of viable oocytes and viable sperm. My estimate for the number of genes necessary for the formation and development of viable sperm is 75. A further 5% (182) of the essential genes are needed make viable progeny, these are the *mel* genes of which I estimate that at least 73 are strictly maternal.

I discovered that two genes (*let-336* and *lin-40*) as wild-type hemizygotes cause increased brood sizes in unmated hermaphrodites. These genes are not only involved in early essential functions but also in the switch from oogenesis to spermatogenesis. *lin-40* also proved to be a complex locus with at least five complementation classes of alleles. I attempted to clone *lin-40* by taking advantage of the fact that, in some cases, there was an extra Tc1 associated with the recessive lethal phenotype. The position obtained on the physical map (A. Coulson and J.E. Sulston pers. comm.) was surprisingly far to the left (near *ges-1*). If this position is correct then there is very little DNA in the gene sparse region of LGV(left), and thus proves that the region has a very high frequency of recombination.

Approximately 70% of the essential genes on LGV(left) have now been identified. It is the largest well characterized genomic region in any higher eukaryote. Because the mutations are easily maintained they have been maintained, and they now are a resource that is available to other researchers. The collection of chromosome breakpoints and transposon tagged genes will facilitate in the correlation of the physical map with the genetic map. Sequencing of the *C. elegans* genome will start shortly (A. Coulson, J.E. Sulston and R.H. Waterston pers. comm.). In conjunction with the sequence information, the phenotypes of the identified genes will provide some of the biology necessary to elucidate the control and function(s) of the genes on LGV(left).

PROPOSALS FOR FUTURE RESEARCH

Screening 30,000 EMS treated chromosomes and mapping the mutations isolated should increase the number of identified genes to 95% of the total. A set of overlapping free duplications might speed the mapping process (method described in Howell *et al.* 1987) The duplications should also allow for the positioning of genes in zone 16. Screening for and mapping additional rearrangements would further sub-divide the zones thus not only easing the mapping of essential genes but also help in the left\right positioning of more of the genes.

The study of the region of high recombination could prove valuable. *lin-40* is putatively cloned and *let-330* is probably tagged with the right break point of *sDf34*. The sequence of these two genes and the region around them would be useful for developing insights into the mechanisms of clustering and the characteristics of genes in regions of high recombination.

In our model of pairing for recombination (Section 3) we postulated that at least three regions contain alignment sites. Finding small deficiencies within these regions, would help localize the alignment sites more precisely. Alternatively, new LGV(left) deficiencies lying outside these regions would either identify more alignment sites or indicate that the deficiencies lie between two such sites.

While mapping lethal mutations in a translocation balanced region, I discovered that a large fraction of the mutations were putative tip deficiencies. I hypothesized that the excision of Tc1 is accompanied by a homolog dependent repair mechanism [similar to the mechanism proposed by Engels *et al.* (1990) for P elements in *D. melanogaster*], but in a translocation balanced region the homolog is not available to effect repair thus producing tip deficiencies. Screening for mobilization of Tc1 induced mutations in a region with and without a translocation balancer would test the hypothesis.

The molecular mapping of Tc1 tagged genes and of selected rearrangement breakpoints would provide a set of markers that anchor the genetic and molecular maps of LGV(left). The anchor points would allow the rapid cloning of genetically identified genes and also for locating molecularly identified genes on the genetic map.

Over 100 essential genes have been identified on LGV(left). The approximate blocking stage has been worked out for most of them. The function(s) of these genes should be identified. A more detailed analysis (through Nomarski optics) of the blocking stages of all the alleles should 1) provide insight into the function(s) of the genes and 2) further define the time of action of the genes. Conditional mutations and suppressors of mutations are also valuable for elucidating gene function. The identification of extragenic suppressor (such as amber suppressors) would provide more tools for dissecting specific gene function. Mosaic analysis (described by Herman 1984) would be useful to assay tissue specific expression.

Appendix A

Alleles of all the genes mapped to zone on LGV(left) are listed. The origins of the alleles are: MH(u60&d11, EMS), TR(d11-u42, EMS), RR(eT1, EMS), GA(eT1, gamma), SP(eT1, spontaneous) all described in Rosenbluth *et al.* (1988); RJ (eT1, EMS screen) Section 4; FM (eT1, formaldehyde screen) Section 1; TC (nT1, Tc1 mobilization screen) Section 2; OT (isolated in other laboratories, but not an exhaustive list of alleles). BLOCK describes the developmental blocking stage of the canonical alleles.

Appendix A (cont)

(alleles of genes on LGV[left])

ZONE	GENE	MH	TR	RR	RJ	GA	SP	OT	FM	TC	Block	Male Rescue
1A	let-450**								s2160		N.D.	
1B1	let-447**				s1457						Maternal	No
									s1654			
	let-336*				s1413						Early	
					s1420							
					s1495							
	let-458*				s1443						Early	
1B2	let-448**									s1363	Mid	
1B3	let-437*				s1405						Mid	
1C	let-453*				s2167						N.D.	
2	let-431**			s1044							Sterile	N.D.
				s1049								
	unc-34							e566				
3	let-x											
4A	let-326			s238								
					s1404						Early	
	unc-60							e677				
					s1586						Early	
								m35				
	ges-1							ca1				
4B	emb-29	s819									Egg	
					s1613							
								g52				
									s1666			
5	let-426	s826										
					s1527						Early	
6A	let-327			s247								
					s1485							
					s1496						Slow	N.D.
	let-478*				s1620						Early	
6B	let-347**			s1035							Late	

Appendix A (cont)

ZONE	GENE	MH	TR	RR	RJ	GA	SP	OT	FM	TC	Block	Male Rescue
7	let-330				s1433							
				s573								
					s1425						Early	
					s1429							
					s1449							
					s1450							
					s1463							
					s1468							
					s1497							
					s1515							
					s1517							
					s1518							
					s1531							
					s1543							
					s1583							
					s1638							
							s1702					
	let-461*				s1486						Early	
8A1	lin-40			s1053								
				s1704								
					s1506						Mid	
					s1593							
					s1611							No
					S1634							No
								e2173				
									s1669			
									s1675			
										s1345		
										s1351		
										s1352		
										s1358		
										s1360		
										s1373		
8A2	let-338					s503					Mid	
				s1020								
	let-455*				s1447						Early	
					s1511							
	let-466*			s990								
				s1063							MEL	N.D.

Appendix A (cont)

ZONE	GENE	MH	TR	RR	RJ	GA	SP	OT	FM	TC	Block	Male Rescue
9A	let-344			s376							Egg	
					s1555							
	let-348				s1436						Early	
					s1448							No
					s1622							
						s998						
	let-430**			s1042							Sterile	N.D.
	let-459*				s1432							
					s1615						Maternal	No
9B	let-341**			s1031							Early	
					s1415							
					s1421							
					s1454							
					s1516							
					s1534							
					s1571							
									s2118			
	let-342**			s1029								
					s1442						Early	
					s1487							
					s1549							
					s1616							
	let-345			s578							Mid	
					s1452							
					s1509							
					s1510							
									s1690			
	unc-62						s472				Egg	
								e644				
10	let-331			s427							Mid(CS)	
	let-350			s250							Sterile	N.D.
									s2126			
	let-415		s129									
					s1505							
					s1525						Early	
	let-417		s204								Early	
				s1313								
					s1424							
									s1679			
	let-419			s219								

Appendix A (cont)

ZONE	GENE	MH	TR	RR	RJ	GA	SP	OT	FM	TC	Block	Male Rescue
10					s1483						Early-mid	
					s1539							
	let-420			s1046							Sterile	
				s1058								
					s1478							No
					s1573							No
					s1584							No
					s1603							No
						s723						
	let-428**			s1070							Sterile	
					s1490							
	let-440*				s1411						Hatches	
					s1440							
					s1552							
					s1560							
					s1589							
	let-452*				s1434						Hatches	
	let-477*				s1608						Slow	N.D.
	let-481*				s1636						Hatches	
	unc-46							e177				
11A1	let-443*				s1417						Early	
11A2	let-401		s193								Mid	
11B1	let-349			s217							Early	
				s572								
						s502						
11B2	let-429			s584							Sterile	
					s1597							No
	let-439*				s1407						Early	
					s1503							
					s1522							
					s1524							
	let-462*				s1481							
					s1590							
					s1594						Early-mid	
	let-473*				s1602						Early	
	let-479*				s1576						Sterile	Yes
11B3	let-329			s575							Early	

Appendix A (cont)

ZONE	GENE	MH	TR	RR	RJ	GA	SP	OT	FM	TC	Block	Male Rescue
11B3	let-463*				s2168						N.D.	
11B4	let-418				s1617						Sterile	No
				s1045								
	let-421				s1460							
				s288								
					s1477						Egg	
					s1632							
11B5	let-422		s194								Early	
				s1312								
					s1548							
					s1563							
					s1578							
						s738						
						s739						
12A	let-402		s127									
					s1526						Early	
						s500						
	let-444*				s1418						Early	
					s1459							
					s1569							Yes
12B	let-403		s120									
				s246								
					s1482						Early	
						s498						
12	unc-83***							e1408				
13	dpy-11							e224				
	let-337	s825										
				s382								
				s1018								
				s1024								
					s1426						Early	
	let-476*				s1621						Sterile	Yes
14	let-410	s815									Mid	
					s1565							
	let-469*				s1582						Sterile	Yes

Appendix A (cont)

ZONE	GENE	MH	TR	RR	RJ	GA	SP	OT	FM	TC	Block	Male Rescue
14	let-471*				s1570						Maternal	Yes
	let-472*				s1605						Mid	
	unc-70		s115									
					s1406							
					s1502							
					s1532							
					s1557						Hatches	
					s1639							
								e524				
15	let-332			s234							Egg	
				s369								
				s1021								
					s1441							
					s1464							
					s1475							
					s1498							
					s1567							
	let-339**				s1444						Early	
					s1469							
				s1019								
	let-343	s816										
				s1025							Egg	
					s1410							
					s1428							
					s1465							
					s1579							
	let-346			s373								
				s1026								
					s1575							
					s1580							
					s1619						Mid	
					s1630							Yes
									s2166			
	let-404		s119								Mid	
	let-425			s385							Sterile	N.D.
	let-438								s2114		N.D.	
	let-442*				s1416						Early-mid	
					s1430							
					s1535							
	let-468*				s1533						Early	
	unc-68							e540				

Appendix A (cont)

ZONE	GENE	MH	TR	RR	RJ	GA	SP	OT	FM	TC	Block	Male Rescue
15	mec-1											
16	let-335			s232								
					s1412							
					s1439						Early	
					s1476							
					s1520							
					s1523							
	let-405	s829										
			s116								Early	
				s388								
	let-406					s514					Mid	
	let-411			s223								
					s1453							
					s1553							
					s1595						Mid	
	let-423	s818									Early	
					s1550							
	let-449**									s1343	Early	
	let-474*				s1577						Early	
	let-480*				s1607						Sterile	No
17	let-408	s827									Egg	
			s195									
	let-413		s128								Egg	
					s1431							
					s1451							
					s1455							
	let-414		s114								Mid	
			s207									
	let-424			s384							Sterile	
				s248								
					s1587							No
	let-436*				s1403						Early	
	let-445*				s1419						Maternal	N.D.
	let-456*				s1479						Early	
18	let-412			s579								
					s1598						Sterile	No
	let-464*				s1504						Early	
					s1530							

Appendix A (cont)

ZONE	GENE	MH	TR	RR	RJ	GA	SP	OT	FM	TC	Block	Male Rescue
18	rol-3			s422								
		s833										
			s126								Early	
				s1030								
				s1040								
					s1408							
					s1409							
					s1473							
					s1494							
					s1519							
						s501						
						s742						
								e754				
19	let-334			s908							Early	
				s383								
	let-340**			s1022								
					s1508						Early	
	let-409	s823										
			s206									
					s1480						Hatch	
					s1507							
					s1528							
					s1546							
					s1547							
	let-416		s113								Late	
20A	let-460*								s1664			
	let-467*				s1521						Hatch	
	unc-23							e25				
20B	let-407	s830										
			s118									
					s1631						Hatch	
									s2122			
	let-441*				s1414						Early	
	let-454*				s1423						Mid	
	let-470*				s1581						Mid	
					s1629							Yes
	let-475*				s1606						Sterile	Yes

* Essential genes identified in this Section.

** Essential genes that I have identified from other screens.

*** *unc-83* was not tested against *sDf37*.

Appendix B

The blocking stage of the alleles of genes on LGV(left) are shown. Some of the results are from Rosenbluth *et al.* (1988).

Appendix B

(blocking stages of essential genes on LGV[left])

GENE	CANONICAL ALLELE	ALLELES	BLOCK	COMMENTS
let-450			N.D.	No alleles are extant
let-447	s1457	s1457	Maternal	Egg let. multi-cell pre-lima bean stage (more eggs over Df)
		s1654	Maternal	Egg let. multi-cell pre-lima bean stage (more eggs over Df)
let-336	s1413	s1413	Early	
		s1420	Early	
		s1495	Early	
let-458	s1443	s1443	Early	
let-448	s1363	s1363	Mid	
let-437	s1405	s1405	Mid	Abnormal tail
let-453	s2167	s2167	N.D.	In same strain as s1424
let-431	s1044	s1044	Ster Adult	
		s1049	Ster Adult	
unc-34	e0566	e0566		
let-x				
let-326		s0238	Mid	
	s1404	s1404	Early	
unc-60	e0677	e0677		
		s1586	Early	
		m0035		
ges-1				
emb-29	s0819	s0819	Egg	
		s1613	Egg	
		g0052	Egg	TS
		s1666	Egg	
let-426		s0826	Mid	
	s1527	s1527	Early	
let-327	s0247	s0247	Slow	Translucent, cold sensitive
		s1485	Slow	Translucent
		s1496	Slow	Morphological abnormalities, lethal at 25c
let-478	s1620	s1620	Early	Possibly 2nd hit, MEL egg lethal over sDf50
let-347	s1035	s1035	Late	
let-330		s1433	Mid	Slow developing as a heterozygote
		s0573		
	s1425	s1425	Early	
		s1429	Early	Occasional mid-larval
		s1449	Mid	
		s1450	Early-Mid	
		s1463	Early	
		s1468	Early	
		s1497	Early-Mid	Ster adult-MEL over sDf34
		s1515	Early	
		s1517	Early	

Appendix B (cont)

GENE	CANONICAL ALLELE	ALLELES	BLOCK	COMMENTS
		s1518	Mid	Occasional slow MEL larval lethals over sDf34
		s1531	Mid	
		s1543	Mid	Hypermorph? Some slow sickly progeny over deficiencies
		s1583	Early-Mid	
		s1638	Mid	MEL larval lethal over let-330(s1433)
		s1702	Mid	
let-461	s1486	s1486	Early	Mid-larval at 15C, there are Sel and Lag alleles
lin-40	e2173	e2173		Vulva blips morphological abnormalities
		s1053	Ster Adult	Vulva blips morphological abnormalities
		s1345		
		s1351		
		s1352	Early-Mid	
		s1358		
		s1360		
		s1373		
		s1506	Mid	Abnormal tail
		s1593	Late	Vulva blips morphological abnormalities
		s1611	Ster Adult	Vulva blips morphological abnormalities
		s1634	Ster Adult	Vulva blips morphological abnormalities
		s1669		Vulva blips morphological abnormalities
		s1675		Vulva blips morphological abnormalities
		s1704	SA/leaky	Vulva blips morphological abnormalities
let-338	s1020	s1020	Mid	
		s0503	Mid	
let-455	s1447	s1447	Early	
		s1511	Early	
let-466		s0990	Maternal	Early larval lethal (F3)
	s1063	s1063	Maternal	Sterile adult (F2)
let-344	s0376	s0376	Egg	
		s1555	Mid	
let-348	s1436	s1436	Early	
		s1448	Maternal	Slow developing MEL egg lethal
		s1622	Early	
		s0998	Mid	
let-430	s1042	s1042	Ster Adult	
let-459	s1432	s1432	N.D.	Possible second mutation in eT1 balanced region
		s1615	Maternal	MEL egg lethal
let-341	s1031	s1031	Earyl	Occasionally blocks as an egg
		s1415	Early	
		s1421	Early	
		s1454	Early	
		s1516	Hatch	
		s1534	Hatch	

Appendix B (cont)

GENE	CANONICAL ALLELE	ALLELES	BLOCK	COMMENTS
		s1571	Early	
		s2118		
let-342		s1029	Mid	Sterile adult over sDf50 (male rescue not tested)
	s1442	s1442	Early	Complements sDf50
		s1487	Early	Male recued sterile adult over sDf50
		s1549	Early	Male recued sterile adult over sDf50
		s1616	Mid	Sterile adult over sDf50 (male rescue not tested), escapers
let-345	s0578	s0578	Mid	
		s1452	Mid	
		s1509	Late	
		s1510	Mid	
		s1690		
unc-62	s0472	s0472	Egg	
		e0644		
let-331	s0427	s0427	Mid(CS)	Slow, cold sensitive
let-350	s0250	s0250	Ster Adult	
		s2126		
let-415		s0129	Late	
		s1505	Late	
	s1525	s1525	Early	
let-417	s0204	s0204	Early	
		s1313		
		s1424	N.D.	In same strain as s2167
		s1679		
let-419		s0219	Mid	
	s1483	s1483	Early-Mid	
		s1539	Early-Mid	
let-420	s1046	s1046	Ster Adult	Vulva blip
		s1058	Ster Adult	Vulva blip
		s1478	Ster Adult	Vulva blip
		s1573	Ster Adult	Vulva Blip
		s1584	Ster Adult	Vulva blip
		s1603	Ster Adult	Vulva blip, abnormal tail
		s0723	Ster Adult	Vulva blip
let-428	s1070	s1070	Ster Adult	
		s1490	Ster Adult	Extra dumpy, extrudes guts
let-440	s1411	s1411	Hatches	
		s1440	Hatches	
		s1552	Htch-Early	
		s1560	Htch-Early	
		s1589	Htch-Early	
let-452	s1434	s1434	Hatches	
let-477	s1608	s1608	Slow	

Appendix B (cont)

GENE	CANONICAL ALLELE	ALLELES	BLOCK	COMMENTS
let-481	s1636	s1636	hatches	
unc-46	e0177			
let-443	s1417	s1417	Early	
let-401	s0193	s0193	Mid	
let-349	s0217	s0217	Early	
		s0572	Late(CS)	Cold sensitive
		s0502	Early	
let-429	s0584	s0584	Ster Adult	
		s1597	Slow	Morphological abnormalities
let-439	s1407	s1407	Early	
		s1503	Early	
		s1522	Early	
		s1524	Early	
let-462		s1481	Mid	Molting problem
		s1590	Mid	
	s1594	s1594	Early-Mid	Blocks earlier over sDf36 than over sDf's 20,26,30 or nDf32
let-473	s1602	s1602	Early	Trapped in old cuticle?
let-479	s1576	s1576	Ster Adult	Lays mushy eggs
let-329	s0575	s0575	Early	
let-463	s2168	s2168	N.D.	In same strain as s1473
let-418		s1617	Ster Adult	Vulva protrudes, partial rescue at 15C
	s1045	s1045	SA/MEL	Vulva protudes @ 24C,
let-421		s1460	Mid	Extra dumpy
		s0288	Leaky	
	s1477	s1477	Egg	
		s1632	Hatch	Extra dumpy, curl, escapers to mid larval (vulva blip)
let-422	s0194	s0194	Early	
		s1312		
		s1548	Early	
		s1563	Early	
		s1578	Mid	Rescued at 15C
		s0738	Early	
		s0739	Early	
let-402		s0127	Mid	
	s1526	s1526	Early	
		s0500	Early	
let-444	s1418	s1418	Early	
		s1459	Mid	
		s1569	Maternal	Maternal egg lethal
let-403		s0120	Mid-Late	
		s0246	Late	
	s1482	s1482	Early	
		s0498	Late	
unc-83	e1408	e1408		
dpy-11	e0224	e0224		
let-337		s0825	Mid	

GENE	CANONICAL ALLELE	ALLELES	BLOCK	COMMENTS
		s0382		
		s1018	Maternal	Maternal egg lethal
		s1024	Maternal	Maternal egg lethal
	s1426	s1426	Early	
let-476	s1621	s1621	Ster Adult	Rescued at 15C, also male rescued
let-410	s0815	s0815	Mid	
		s1565	Mid	
let-469	s1582	s1582	Ster Adult	
let-471	s1570	s1570	Maternal	MEL larval lethal,
let-472	s1605	s1605	Mid	Morphological abnormalities
unc-70		s0115		Most unc-70 lethal alleles are slightly
		s1406	Early	Uncoordinated as heterozygotes
		s1502	Hatches	Weak coiler or folded; and escapers to mid larval
		s1532	Early	weak coiler or folded
		s1557	Hatches	
		s1639	Hatches	Coiled or folded
	e0524	e0524		
let-332	s0234	s0234	Egg	
		s0369	Egg	
		s1021	Early	Escapers
		s1441	Egg	
		s1464	Egg	Well developed worms in eggs
		s1475	N.D.	
		s1498	Egg	
		s1567	Egg	
let-339	s1444	s1444	Early	
		s1469	Early-Mid	Some eggs do not hatch
		s1019	Maternal	Cold sensitive
let-343		s0816	Early-Mid	Maternal cold sensitive
	s1025	s1025	Egg	
		s1410	Early	
		s1428	Early-Mid	
		s1465	Slow	Cold sensitive
		s1579	Early	
let-346		s0373	Late	
		s1026	Late	
		s1575	Mid-Late	
		s1580	Mid-Late	Most have protruding vulvae
	s1619	s1619	Mid	Possibly reflexes backwards
		s1630	Ster Adult	Possibly reflexes backwards
		s2166		
let-404	s0119	s0119	Mid	
let-425	s0385	s0385	Ster Adult	

Appendix B (cont)

GENE	CANONICAL ALLELE	ALLELES	BLOCK	COMMENTS
let-438	s2114	s2114		
let-442	s1416	s1416	Early-Mid	
		s1430	Early-Mid	Partial rescue at 15C (Sterile Adults)
		s1535	Early-Mid	
let-468	s1533	s1533	Early	
unc-68	e0540			
mec-1				
let-335		s0232	Mid	
		s1412	Late	
	s1439	s1439	Early	
		s1476	Early	Tends to coil
		s1520	Early	Escapers (Sterile adults with vulva blips)
		s1523	Early	
let-405		s0829	Mid	
	s0116	s0116	Early	
		s0388	Mid	
let-406	s0514	s0514	Mid	
let-411		s0223	Late	
		s1453	Mid	
		s1553	Mid	
	s1595	s1595	Mid	
let-423	s0818	s0818	Early	
		s1550	Hatch/Esc	Hatch, Early, Mid Blockers
let-449	s1343	s1343	Early	
let-474	s1577	s1577	Early	
let-480	s1607	s1607	Ster Adult	
let-408	s0027	s0827	Egg	
		s0195	Late	
let-413	s0128	s0128	Egg	
		s1431	Egg	
		s1451	Egg	
		s1455	Early	
let-414	s0114	s0114	Mid	
		s0207	Mid	
let-424	s0384	s0384	Ster Adult	
		s0248	Ster Adult	
		s1587	Ster Adult	
let-436	s1403	s1403	Early	
let-445	s1419	s1419	Maternal	Tight coilers; MEL larval lethal
let-456	s1479	s1479	Early	
let-412		s0579	Maternal	Maternal egg lethal
	s1598	s1598	Ster Adult	Roller

GENE	CANONICAL ALLELE	ALLELES	BLOCK	COMMENTS
let-464	s1504	s1504	Early	Approx. 1/2 do not hatch
		s1530	Early	
rol-3		s0422	Mid	
		s0833	Mid	
	s0126	s0126	Early	
		s1030		
		s1040	Mid (20C)	Temperature sensitive
		s1408	Early	
		s1409	Early	
		s1473	N.D.	In same strain as s2168
		s1494	Early	
		s1519	Early	
		s0501		
		s0742		
		e0754		
let-334	s0908	s0908	Early	
		s0383	Mid	
let-340		s1022	Mid	
	s1508	s1508	Early	
let-409		s0823	Early	
		s0206	Early	
	s1480	s1480	Hatch	Tend to curl
		s1507	Hatch	
		s1528	Hatch	
		s1546	Early	
		s1547	Hatch	
let-416	s0113	s0113	Late	
unc-23	e0025	e0025		
let-460	s1664	s1664		
let-467	s1521	s1521	Hatch	
let-407		s0830	Early	
		s0118	Early	
	s1631	s1631	Hatch	
		s2122		
let-441	s1414	s1414	Early	Escapers (not due to recomb.), rescued at 15C
let-454	s1423	s1423	Mid	Double cuticle, no pumping, probably stuck in molt
let-470	s1581	s1581	Mid	
		s1629	Slow	Rescued at 15C
let-475	s1606	s1606	Ster Adult	Rescued at 15C

REFERENCES

Auerbach, C., M. Moutschen-Dahmen and J. Moutschen (1977) Genetic and cytogenetical effects of formaldehyde and related compounds. Mutation Res. 39: 317-362.

Baker, B.S., A.T.C. Carpenter, M.S. Esposito and L. Sandler (1976) The genetic control of meiosis. Annu. Rev. Genet. 10: 53-134.

Barnett, T., C. Pachel, J.P. Gergen, and P.C. Wensink (1980) The isolation and characterization of Drosophila yolk protein genes. Cell 21: 729-738.

Barton, M.K. and J. Kimble (1990) *fog-1*, a regulatory gene required for specification of spermatogenesis in the germ line of *Caenorhabditis elegans*. Genetics 125: 29-39.

Barton, M.K., T. Schedl and J. Kimble (1987) Gain-of-function mutations of *fem-3*, a sex-determination gene in *Caenorhabditis elegans*. Genetics 115: 107-119.

Bejsovec, A., and P. Anderson (1988) Myosin heavy chain mutations that disrupt *C. elegans* thick filament assembly. Genes Dev. 2: 1307-1317.

Belote, J.M., F.M. Hoffman, M. McKeown, R.L. Chorsky and B.S. Baker (1990) Cytogenetic analysis of chromosome region 73AD of *Drosophila melanogaster*. Genetics 125: 783-793.

Benian, G.M., J.E. Kiff, N. Neckelmann, D.G. Moerman and R.H. Waterston (1989) Sequence of an unusually large protein implicated in regulation of myosin activity in *C. elegans*. Nature 342: 45-50.

Brenner, S. (1974) The genetics of *Caenorhabditis elegans*. Genetics 77: 71-94.

Brown, S.J. (1984) Genetic interactions affecting muscle organization in the nematode *C. elegans*. Ph.D. Thesis Univ. of Missouri, Columbia.

Chovnick, A., V. Finnerty, A. Schalet and P. Duck (1969) Studies on genetic organization in higher organisms. I. Analysis of a complex gene in *Drosophila melanogaster*. Genetics 62: 145-160.

Clark, D.V., (1990) The *Unc-22(IV)* of *Caenorhabditis elegans:* Genetic analysis and molecular mapping. Ph.D Thesis, Simon Fraser Univ., Burnaby B.C., Canada.

Clark, D.V., R.C. Johnsen K.S. McKim and D.L. Baillie (1990) Analysis of lethal mutations induced in a mutator strain that activates transposable elements in *Caenorhabditis elegans*. Genome 33-1: 109-114.

Clark, D.V., T.M. Rogalski, L.M. Donati and D.L. Baillie 1988. The *unc-22(IV)* region of *Caenorhabditis elegans*: Genetic analysis of lethal mutations. Genetics 119: 345-353.

Collins, J., B. Saari and P. Anderson (1987) Activation of a transposable element in the germ line but not in the soma of *Caenorhabditis elegans*. Nature 328: 726-728.

Coulson, A., J. Sulston, S. Brenner and J. Karn (1986) Towards a physical map of the genome of the nematode *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA 83: 7821-7825.

Coulson, A., R. Waterston, J. Kiff, J. Sulston, and Y. Kohara (1988) Genome linking with yeast artificial chromosomes. Nature 335: 184-186.

Crain, W.R., F.C. Eden, W.R. Pearson, E.H. Davidson and R.J. Britten (1976) Absence of short period interspersion of repetitive and non-repetitive sequences in the DNA of *Drosophila melanogaster*. Chromosoma 56: 309-26.

Crow, E.L. and Gardner (1959) Confidence intervals for the expectation of a Poisson variable. Biometrika 46: 441-453.

Davis, R.W., D. Botstein and J.R. Roth (1980) A manual for genetic engineering: Advanced bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Delong, L., L.P. Casson and B.J. Meyer (1987) Assessment of X chromosome dosage compensation in *Caenorhibditis elegans* by phenotypic analysis of *lin-14*. Genetics 117: 657-670.

Dibb, N.J., D.M. Brown, J. Karn, D.G. Moerman, S.L. Bolten and R.H. Waterton (1985) Sequence analysis of mutations that affect the synthesis, assembly and enzymatic activity of the *unc-54* myosin heavy chain of *Caenorhabditis elegans*. J. Mol. Biol. 183: 543-551.

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 Univ., Burnaby B.C., Canada.

Doniach, T. (1986) Genetic analysis of sex determination in the nematode *Caenorhabditis elegans*. Ph.D. thesis, Council of National Academy Awards, United Kingdom.

Edgley, M.L. and D.L. Riddle (1990) The nematode *Caenorhabditis elegans*. in: *Genetic Maps Locus maps of complex genomes*. Fifth ed. CSH Laboratories Edited by S.J. O'Brien. 3.111-3.133. Cold Spring Harbor, New York.

Eide, D., and P. Anderson (1985) Transposition of Tc1 in the nematode *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA 82: 1756-1760.

Eide, D., and P. Anderson (1988) Insertion and excision of *Caenorhabditis elegans* transposable element Tc1. Molec. Cell. Biol. 8: 737-746.

Emmons, S.W., M.R. Klass and D. Hirsh (1979) Analysis of the constancy of DNA sequences during development and evolution of the nematode *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA 76-3: 1333-1337.

Emmons, S.W., L. Yesner, K. Ruan and D. Katzenberg (1983) Evidence for a transposon in *Caenorhabditis elegans*. Cell 32: 55-65.

Engels, W.R., D.M. Johnson-Schlitz, W.B. Eggleston and J. Sved (1990) High-frequency P element loss in *Drosophila* is homolog-dependent. Cell 62: 515-525.

Feldman, M.Ya. (1975) Reactions of nucleic acids and nucleoproteins with formaldehyde. Progr. in Nuçleic Acid Res. Mol. Biol. 13: 1-49.

Feinberg, A. and B. Vogelstein (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132: 6-13.

Ferguson, E.L. and H.R. Horvitz (1985) Identification and characterization of 22 genes that affect vulva cell lineages of the nematode *Caenorhabditis elegans*. Genetics 110: 17-72.

Fincham, J.R.S. (1983) Genetics. Jones and Bartlett Publishers, Inc. Boston, USA.

Finnerty, V. and A. Chovnick (1970) Studies on genetic organization in higher organisms III. Confirmation of the single cistron-allele complementation model of organization of the *maroon-like* region of *Drosophila melanogaster*. Genet. Res. Camb. 15: 351-355.

Finnerty, V., P. Duck and A. Chovnick (1970) Studies on genetic organization in higher organisms, II. Complementation and fine structure of the *maroon-like* locus of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U.S.A. 65: 939-946.

Gertler, F.B., J.S. Doctor and F.M. Hoffmann (1990) Genetic suppression of mutations in the Drosophila abl proto-oncogene homolog. Science 248: 857-860.

Geyer, P.K., K.L. Richardson, V.G. Corces and M.M. Green (1988) Genetic instability in *Drosophila melanogaster*: P-element mutagenesis by gene conversion. Proc. Natl. Acad. Sci. U.S.A. 85: 6455-6459.

Greenwald, I. (1985) *lin-12*, a nematode homeotic gene, is homologous to a set of mammalian proteins that includes epidermal growth factor. Cell 43: 583-590.

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 Research 15: 2295-2307.

Hadorn, E. 1945 "Zur pleiotropie der genwirkung." Arch. Jul. Klaus-Stiftg., Erganzungsband zu Band 20: 82-95.

Hadorn, E. (Translated by U. Mittwoch) 1961 Developmental genetics and lethal factors. Metheun & Co. Ltd. London England.

Hawley, R.S. (1980) Chromosomal sites necessary for normal levels of meiotic recombination in *Drosophila melanogaster*. I. Evidence for mapping of the sites. Genetics 94: 625-646.

Herman, R.K. (1984) Analysis of genetic mosaics of the nematode *Caenorhabditis elegans*. Genetics 88: 49-65.

Herman, R.K. and C.K. Kari (1989) Recombination between small X chromosome duplications and the X chromosome in *Caenorhabditis elegans*. Genetics 121: 723-737.

Herman, R.K., C.K. Kari and P.S. Hartman (1982) Dominant X-chromosome non-disjunction mutants of *Caenorhabditis elegans*. Genetics 102: 379-400.

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.

Hilliker, A.J., S.H. Clark, A. Chovnick and W.M. Gelbart (1980) Cytogenetic analysis of the chromosomal region immediately adjacent to the *rosy* locus in *Drosophila melanogaster*. Genetics 95: 95-110.

Hochman, B. (1970) Analysis of chromosome 4 in *Drosophila melanogaster*. II: ethyl methanesulfonate induced lethals. Genetics 67: 235-252.

Hodgkin, J. (1983a) X chromosome dosage and gene expression in *Caenorhabditis elegans*: two unusual dumpy genes. Mol. Gen Genet. 192: 452-458.

Hodgkin, J. (1983b) Two types of sex determination in a nematode. Nature 304: 267-268.

Hodgkin, J. (1985) Novel nematode amber suppessors. Genetics 111: 287-310.

Hodgkin, J. (1986) Sex determination in the nematode *C. elegans*: analysis of *tra-3* suppressors and characterization of *fem* genes. Genetics 114: 15-52.

Hodgkin, J. (1987) Sex determination and dosage compensation in *Caenorhabditis elegans*. Annu. Rev. Genet. 21: 133-154.

Hodgkin, J. (1988) The nematode *Caenorhabditis elegans*. Wood, W.B. *et al.* (eds) Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. 243-279.

Hodgkin, J. (1990) Sex determination compared in *Drosophila* and *Caenorhabditis*. Nature 344: 721-728.

Hodgkin, J. and S. Brenner (1977) Mutations causing transformation of sexual phenotype in the nematode *Caenorhabditis elegans*. Genetics 86: 275-287.

Hodgkin, J., H.R. Horvitz and S. Brenner (1979) Nondisjunction mutants of the nematode *Caenorhabditis elegans*. Genetics 91: 67-94.

Horvitz, H.R., S. Brenner, J. Hodgkins and R.K. Herman (1979) A uniform genetic nomenclature for the nematode *Caenorhabditis elegans*. Mol. Gen. Genet. 175: 129-133.

Howell, A.M. (1989) Essential genes in a region of chromosome I in *Caenorhabditis elegans*. Ph.D. Thesis, University of British Columbia, Vancouver, B.C. Canada.

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.

Howell, A.M. and A.M. Rose (1990) Analysis of essential genes in the *hDf6* region of chromosome I in *Caenorhabditis elegans*. Genetics (In press).

Johnsen, R.C. and D.L. Baillie (1988) Formaldehyde mutagenesis of the *eT1* balanced region in *Caenorhabditis elegans*: Dose-response curve and the analysis of mutational events. Mutation Res. 201: 137-147.

Johnsen, R.C., R.E. Rosenbluth and D.L. Baillie (1986) Genetic analysis of linkage group V(left) in *Caenorhabditis elegans*. Genetics 113: s11.

Judd, B.H., M.W. Shen and T.C. Kaufman (1972) The anatomy and function of a segment of the X-chromosme of *Drosophila melanogaster*. Genetics 71: 139-156.

Kaplan, W.D. (1948) Formaldehyde as a mutagen in *Drosophila*. Science 108: 43.

Karlik, C.C., J.W. Mahaffey, M.D. Coutu and E.A. Fyrberg (1984) Organization of contractile protein genes within the 88F subdivision of the *Drosophila melanogaster* third chromosome. Cell 37: 469-481.

Kemphues, K.J., M. Kusch and N. Wolf (1988) Maternal-effct lethal mutations on linkage group II of *Caenorhabditis elegans*. Genetics 120: 977-986.

Kim, J.S. and A.M. Rose (1987) The effect of gamma radiation on recombination frequency in *Caenorhabditis elegans*. Genome 29: 457-462.

Kimble, J.E. and S. Ward (1988) The nematode *Caenorhabditis elegans*. Wood, W.B. *et al.* (eds) Cold Spring Harbor, New York. 243-279.

Kimble, J.E. and J.G. White (1981) On the control of germ cell development in *Caenorhabditis elegans*. Dev. Biol. 81: 208-219.

Lambie, E.J. and G.S. Roeder (1986) Repression of meiotic crossing over by a centromere (CEN3) in *Saccharomyces cerevisiae*. Genetics 114: 769-789

Lefevre, G. and W. Watkins (1986) The question of the total gene number in *Drosophila melanogaster*. Genetics 113: 869-895.

L'Hernault, S.W., D.C. Shakes and S. Ward (1988) Developmental genetics of chromosome I spermatogenesis-defective mutants in the nematode *Caenorhabditis elegans*. Genetics 120: 435-452.

MacLeod, A.R., R.H. Waterston, R.M. Fishpool and S. Brenner (1977) Identification of the structural gene for a myosin heavy-chain in *Caenorhabditis elegans*. Mol. Biol. 114: 133-140.

Madl, J.E. and R.K. Herman (1979) Polyploids and sex determination in *Caenorhabditis elegans*. Genetics 93: 393-402.

Mains, P.E., I.A. Sulston and W.B. Wood (1990) Dominant maternal-effect mutations causing embryonic lethality in *Caenorhabditis elegans*. Genetics 125:351-369.

Maniatis, T., E.F. Fritsch and J. Sambrook (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Martin, C.H. and E.M. Meyerowitz (1988) Mosaic evolution in the Drosophila genome. Bioessays 9: 65-69.

McCoubrey, W.K., K.D. Nordstrom and P.M. Meneely (1988) Microinjected DNA from the X chromosome affects sex determination in *Caenorhabditis elegans*. Science 242: 1146-1151.

McKim, K.S., M.F.P. Heschl, R.E. Rosenbluth and D.L. Baillie (1988a) Genetic organization of the *unc-60* region in *Caenorhabditis elegans*. Genetics 118: 49-59.

McKim, K.S., A.M. Howell and A.M. Rose (1988b) The effects of translocations on recombination frequency in *Caenorhabditis elegans*. Genetics 120: 987-1001.

McKim, K.S. and A.M. Rose (1990) Chromosome I duplications in *Caenorhabditis elegans*. Genetics 124: 115-132.

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.

Meneely, P.M. and W.B. Wood (1987) Genetic analysis of X-chromosme dosage compensation in *Caenorhabditis elegans*. Genetics 117: 25-41.

Meyer, B.J. and L.P. Casson (1986) *Caenorhabditis elegans* compensates for the difference in X chromosome dosage between the sexes by regulating transcript levels. Cell 47: 871-881.

Miller, L.M., J.D. Plenefisch, L.P. Casson and B.J. Meyer (1988) *xol-1*: A gene that controls the male modes of both sex determination and X chromosome dosage compensation in *C. elegans*. Cell 55: 167-183.

Moerman, D.G. and D.L. Baillie (1981) Formaldehyde mutagenesis in *Caenorhabditis elegans*. Mutation Res. 80: 173-279.

Moerman, D.G., G.M. Benian, R.J. Barstead, L.A. Schriefer and R.H. Waterston (1988) Identification and intracellular localization of the *unc-22* gene product of *Caenorhabditis elegans*. Genes and Development 2: 93-105.

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.

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. and R.H. Waterston (1989) Mobile elements in *Caenorhabditis elegans* and other nematodes. In *Mobile DNA* edited by D. Berg and M. Howe, American Society of Microbiology, Washington D.C. pp 537-556.

Mori, I., G.M. Benian, D.G. Moerman and R.H. Waterston (1988a) Transposable element Tc1 of *Caenorhabditis elegans* recognizes specific target sequences for integration. Proc. Natl. Acad. Sci. USA 85: 861-864.

Mori, I., D.G. Moerman and R.H. Waterston (1988b) Analysis of a mutator activity necessary for germline transposition and excision of Tc1 transposable elements in *Caenorhabditis elegans*. Genetics 120: 397-407.

Muller, H.J. (1932) Further studies on the nature and causes of gene mutations. Proceedings of the Sixth International Congress of Genetics 1: 213-255.

Nelson, G.A., W.W. Schbert, T.M. Marshall, E.R. Benton and E.V. Benton (1989) Radiation effects in *Caenorhabditis elegans*: mutagenesis by high and low LET ionizing radiation. Mutation Res. 212: 181-192.

Nigon, V. (1949) Les modalites de la reproduction et le determinisme du sexe chez quelques nematodes libres. Ann. Sci. Nat. (Zool.) 11: 1-132.

Nusbaum, C. and B.J. Meyer (1989) The *Caenorhabditis elegans* gene *sdc-2* controls sex determination and dosage compensation in XX animals. Genetics 122: 579-593.

O'Donnell, J., H.C. Mandel, M. Krauss and W. Sofer (1977) Genetic and cytogenetic analysis of the *Adh* region in *Drosophila melanogaster*. Genetics 86: 553-566.

Park, E.-C. and H.R. Horvitz (1986a) Mutations with dominant effects on the behavior and morphology of the nematode *Caenorhabditis elegans*. Genetics 113: 821-852.

Park, E.-C. and H.R. Horvitz (1986b) *C. elegans unc-105* mutations affect muscle and are suppressed by other mutations that affect muscle. Genetics 113: 853-867.

Peacock, W.J., D. Brutlag, E. Goldring, R. Appels, C.W. Hinton and D.L. Lindsley (1974) The organization of highly repeated DNA sequences in *Drosophila melanogaster* chromosomes. Cold Spring Harbor Symposia on Quantitative Biology 38: 405-416.

Prasad, S.S. and D.L. Baillie (1989) Evolutionarily conserved coding sequences in the *dpy-20 - unc-22* region of *Caenorhabditis elegans*. Genomics 5: 185-198.

Priess, J.R., H. Schnabel and R. Schnabel (1987) The *glp-1* locus and cellular interactions in early *Caenorhabditis elegans* embryos. Cell 51: 601-611.

Rand, J.B. (1989) Genetic analysis of the *cha-1-unc-17* gene complex in *Caenorhabditis*. Genetics 122: 73-80.

Rapoport, I.A. (1946) Carbonyl compounds and the chemical mechanisms of mutation. Compt. Rend. Acad. Sci. (U.R.S.S.) 54: 65-67.

Rattray, B. and A.M. Rose (1988) Increased intragenic recombination and non-disjunction in the *rec-1* strain of *Caenorhabditis elegans*. Genet. Res. 51: 89-93

Rigby, P.W., M. Dieckmann, C. Rhodes and P. Berg (1977) Labelling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase. J. Mol. Biol. 113: 237-251.

Rogalski, T.M. and D.L. Baillie (1985) Genetic organization of the *unc-22* IV region of *Caenorhabditis elegans*. Mol. Gen. Genet. 201: 409-414.

Rogalski, T.M., A.E. Bullerjahn and D.L. Riddle (1988) Lethal and amanitin-resistance mutations in the *Caenorhabditis elegans ama-1* and *ama-2* genes. Genetics 120: 409-422.

Rogalski, T.M., D.G. Moerman and D.L. Baillie (1982) Essential genes and deficiencies in the *unc-22* region of *Caenorhabditis elegans*. Genetics 102: 725-736

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.

Rose, A.M. and D.L. Baillie (1979a) Effect of temperature and parental age on recombination and nondisjuction in *Caenorhabditis elegans*. Genetics 92: 409-418.

Rose, A.M. and D.L. Baillie (1979b) A mutation in *Caenorhabditis elegans* that increases recombination frequency more than threefold. Nature 281: 599-600.

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.

Rose, A.M. and K.S. McKim (1989) Meiotic recombination in *Caenorhabditis elegans*. in *P. and S. Biomedical Sciences Symposia*, edited by M.E. Gottesman and H.J. Vogel. Academic Press, New York(in press)

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., C. Cuddeford and D.L. Baillie (1983) Mutagenesis in *Caenorhabditis elegans* I. A rapid eukaryotic mutagen test system using the reciprocal translocation, *eT1(III,V)*. Mutation Res. 110: 39-48.

Rosenbluth, R.E., C. Cuddeford and D.L. Baillie (1985) Mutagenesis in *Caenorhabditis elegans* II. A spectrum of mutational events induced with 1500 R of gamma-radiation. Genetics 109: 493-511.

Rosenbluth, R.E., R.C. Johnsen and D.L. Baillie (1990) Pairing for recombination in LGV of *Caenorhabditis elegans*: A model based on recombination in deficiency heterozygotes. Genetics 124: 615-625.

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. 52: 105-118.

Schedl, T. and J. Kimble (1988) *fog-2*, a germ-line-specific sex determination gene required for hermaphrodite spermatogenesis in *Caenorhabditis elegans*. Genetics 119: 43-61.

Schmid, C.W., J.E. Manning and N. Davidson (1975) Inverted repeat sequences in the *Drosophila* genome. Cell 5: 159-172.

Shannon, M.P., T.C. Kaufman, M.W. Shen and B.H. Judd (1972) Lethality patterns and morphology of selected lethal and semi-lethal mutations in the zeste-white region of *Drosophila melanogaster*. Genetics 72: 615-638.

Sigurdson, D.C., G.J. Spanier and R.K. Herman (1984) *Caenorhabditis elegans* deficiency mapping. Genetics 108: 331-345.

Simmons, M.J. and J.K. Lim (1980) Site specificity of mutations arising in hybrids of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U.S.A. 77: 6042-6046.

Simmons, M.J., J.D. Raymond, N.A. Johnson and T.M. Fahey (1984) A comparison of mutation rates for specific loci and chromosome regions in dysgenic hybrid males of *Drosophila melanogaster*. Genetics 106: 85-94.

Slizynska, H. (1957) Cytological analysis of formaldehyde induced chromosomal changes in *Drosophila melanogaster*. Proc. Roy. Soc. Edin. 66B: 228-304.

Snutch, T.P. (1984) A molecular and genetic analysis of the heat shock response of *Caenorhabditis elegans*. Ph.D. Thesis, Simon Fraser University, Burnaby, B.C., Canada.

Starr, T., A.M. Howell, J. McDowall, K. Peters and A.M. Rose (1989) Isolation and mapping of DNA probes within the linkage group I gene cluster of *Caenorhabditis elegans*. Genome 32: 365-372.

Stewart, H.I., R.E. Rosenbluth and D.L. Baillie (1990) Most ultraviolet irradiation induced mutations in the nematode *Caenorhabditis elegans* are chromosome rearrangements. Mutation Res. (submitted).

Sulston, J.E. and H.R. Horvitz (1977) Post-embryonic cell lineages of the nematode *C. elegans*. Dev. Biol. 56: 110-156.

Sulston, J.E., E. Schierenberg, J.G. White and J.N. Thomson (1983) The embryonic cell lineages of the nematode *C. elegans*. Dev. Biol. 110: 64-119.

Suzuki, D.T., L.K. Piternick, S. Hayashi, M. Tarasoff. D. Baillie and U. Erasmus (1967) Temperature-sensitive mutations in *Drosophila melanogaster*, I. Relative frequencies among gamma-ray and chemical induced sex-linked recessive lethals and semilethals. Proc. Natl. Acad. Sci. USA 57: 907-912.

Szauter, P. (1984) An analysis of regional constraits on exchange in *Drosophila melanogaster* using recombination-defective meiotic mutants. Genetics 106: 45-71.

Villeneuve, A.M. and B.J. Meyer (1987) *sdc-1*: A link between sex determination and dosage compensation in *C. elegans*. Cell 48: 25-37.

von Ehrenstein, G. and E. Schierenburg (1979) Cell lineages and development of *Caenorhabditis elegans* and other nematodes. in Nematodes as model biological systems, B.M. Zuckerman (ed.) Academic Press, New York.

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

White, J. (1988) The nematode *Caenorhabditis elegans*. Wood, W.B. *et al.* (eds) Cold Spring Harbor, New York. 243-279.

Wright, T.R.F. (1987) The genetic and molecular organization of the dense cluster of functionally related, vital genes in the DOPA decarboxylase region of the *Drosophila melanogaster* genome. In: Structure and function of eukaryotic chromosomes. Edited by W. Hennig. Springer Verlag, Berlin.

Wright, T.R.F., E.Y. Wright, B.C. Black, E.S. Pentz, C.P. Bishop, J. Kullman, M.H. Corjay and G.R. Hankins (1984) The genetic and molecular organization of a gene cluster involved in catecholamine metabolism and sclerotisation. Crete Drosophila Meeting, Kolymbari (Abst.)

Yochem, J. and I. Greenwald (1989) *glp-1* and *lin-12*, genes implicated in distinct cell-cell interactions in *C. elegans*, encode similar transmembrane proteins. Cell 58: 553-563.