A GENETIC CHARACTERIZATION OF THE REGION AROUND THE UNC-22 GENE IN CAENORHABDITIS ELEGANS

by

Teresa Maria Rogalski

B.Sc. (Hons), Simon Fraser University, 1979

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, 1983

APPROVAL

Name: TERESA M. ROGALSKI

Degree: Doctor of Philosophy

Title of Thesis: A genetic characterization of the region around the unc-22 IV gene in Caenorhabditis elegans

Examining Committee

Chairman: Dr. Glen H. Geen

Dr. David L. Baillie, Senior Supervisor

Dr. Andrew T. Beckenbach

Dr. B. M. Honda

Dr. C. L. Kemp

Dr. Michael J. Smith

Dr. J. M. Webster

Dr. David G. Holm, Assoc. Professor, Dept. of Zoology, UBC, Public Examiner

Dr. M. Mackauer, Professor, Dept. of Biological Sciences, SFU, Public Examiner

Dr. Arthur Chovnik, Professor of Biology, Genetics & Cell Biology Section, Biological Sciences, University of Connecticut, External Examiner.

Date approved Oct 20, 1983

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

A genetic characterization of the region around the unc-22 IV Caenorhabditis elegans

Author: _____
(signature)

Teresa M. Rogalski
(name)

November 24, 1983
(date)

ABSTRACT

The aim of this analysis was to establish a genetically well defined region of the genome of Caenorhabditis elegans in order to study the organization of the genetic material in a higher eucaryote. The sDf2 region described in this thesis is the most extensively characterized region in this organism. Twenty essential and two nonessential genes have been identified in an interval of 1.5-map units on Linkage Group IV. The mutations which define these genes were positioned by two independent methods: recombination mapping and complementation mapping with the deficiencies, sDf7, sDf8, sDf9 and sDf10. With few exceptions, the positions obtained by these two mapping methods were in very good agreement. Eight of the twenty essential genes have more than one allele, indicating that these genes represent about two-thirds of the essential genes in this region. Therefore, assuming a total of thirty-two essential and two nonessential genes, the genes in this region are separated by approximately 0.05-map units.

The genetic analysis of the sDf2 region was done in conjunction with a fine-structure analysis of unc-22 IV, a muscle gene in this region. The maximum size of the unc-22 gene, which is defined as the distance between the adjacent genes on either side of unc-22 , is 0.1-map units while the minimum size of this gene, which is represented by the fine-structure map, is 0.02-map units. At present, the unc-22 coding element includes the right end of the fine-structure map. Three possible internal

deletions of unc-22 have been identified and positioned. In addition, the right endpoint of a deficiency or an inversion affecting the adjacent genes, let-56 and unc-22, has been positioned inside the unc-22 gene. This formaldehyde-induced mutation is associated with a dominant Unc-22 phenotype indicating that the mutant unc-22 gene is transcribed. The simplest interpretation of these data is that the left end of the unc-22 gene has been deleted and this gene is, therefore, transcribed from right to left on the genetic map.

Attempts to clone the unc-22 gene and the adjacent region are in progress in order to compare the genetic map of this region with its DNA sequence.

ACKNOWLEDGEMENTS

I would like to thank my senior supervisor, Dr. D.L.Baillie, my colleague, R.E.Rosenbluth, and the members of my supervisory committee, Dr. A.T. Beckenbach, Dr. B.A.Honda, Dr. C.L.Kemp, Dr. M.J.Smith and Dr. J.M.Webster, for their help and encouragement during the past four years; my colleagues, Dr. D.G.Moerman, Dr. A.M.Rose, T.P.Snutch, D.Nelson, S.S.Prasad and C.Cuddeford for their conversation and suggstions; R.E.Rosenbluth and Dr. D.L.Baillie for their help in the fine-structure screens; Dr.D.L.Baillie, R.E.Rosenbluth and C.Cuddeford for their help in the lethal screen; D.R.Wilson for his help with the computer; M.M.Rogalski for her help in typing this thesis; and C.Salamanca, J.Gomez, H.Hemani, L.Hale and M.M.Rogalski for their technical assistance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

A. INTRODUCTION

In contrast to procaryotes, most eucaryotes are complex multicellular organisms composed of many different cell types. Generally, these various cell types all contain the same DNA sequences or genes, but are different from each other because they synthesize and accumulate different sets of proteins (Gurdon 1962; Garrels 1979). The development of a multicellular eucaryote from a single cell is achieved by regulating the expression of different genes or sets of genes in certain cells at specific times during development. The molecular mechanisms which regulate gene expression in eucaryotes are unknown.

Genomic DNA in eucaryotes is bound to histone proteins to form a histone-DNA complex (chromatin) which is organized into a series of looped domains and undergoes several orders of folding to form the chromosomes (reviewed by Isenberg 1979; Alberts et al. 1983). Other, nonhistone proteins are also associated with eucaryotic DNA. Some of these, presumably, are important for maintaining the highly condensed state of this DNA. Most eucaryotic genes contain intervening sequences (introns) which are transcribed but not translated. Transcription and RNA-processing occur in the nucleus where the chromosomes are located. Processing of the primary RNA transcript to the mature mRNA involves the capping of the 5′ end, the addition of poly A to the 3′ end (most mRNAs) and the removal or splicing out of

introns (Perry 1981; Chambon 1981). After processing, the mature mRNA is transported to the cytoplasm where translation occurs. The regulation of gene expression in eucaryotes could occur at any one of the steps described above. Although post-transcriptional (Early et al. 1980) and translational (Gross 1967) control of gene expression do occur, the primary form of regulation seems to operate at the level of transcription (Tsai et al. 1979; Derman et al. 1981). Genes which are turned on in a cell are transcribed and their mRNA is translated whereas genes which are turned off in a cell are not transcribed.

Genes in procaryotes are also regulated at the level of transcription. The molecular mechanisms of regulation have been worked out for several bacterial operons (Miller and Reznikoff 1978) and for the lysis-lysogeny decision in phage lambda (Ptashne et al. 1980; Herskowitz and Hagen 1980). In most cases, the gene or operon is regulated by a protein or proteins which bind to DNA sequences 5′ to the structural elements which they control to either activate or repress transcription. Similar mechanisms may exist in higher eucaryotes.

In order to understand the mechanisms of eucaryotic gene activation it is necessary to understand how these genes are organized in the genome. Eucaryotic genomic organization has been intensively studied in Drosophila melanogaster by both genetic and molecular techniques (reviewed by Spradling and Rubin 1981). The haploid genome of D. melanogaster contains

approximately 1.65 x 10^8 base pairs of DNA (Rasch et al. 1971). About 70% of this DNA is single copy while the remaining 30% is composed of repeated sequences. The repetitive class includes the highly repetitive centromeric and telomeric sequences, the genes coding for rRNA, 5S RNA and histones which are tandemly repeated in the genome, the tRNA genes which are clustered in about 30 chromosomal regions and members of multigene protein families. The majority of the repetitive DNA in D. melanogaster is composed of about 40 families of transposable elements dispersed throughout the genome. The single copy DNA (about 1.2 x 108 bp) includes most of the genes that code for proteins and presumably their regulatory elements or sequences.

The best estimate of the total number of genes in D. melanogaster comes from genetic analyses of defined chromosomal regions. The number of genes or complementation groups in several regions have been determined by saturation mapping (Judd et al. 1972; O'Donnell et al. 1977; Young and Judd 1978; Hochman 1978; Woodruff and Ashburner 1979; Hilliker et al. 1980; Gausz et al. 1979; Belyaeva et al. 1980; Wright et al. 1981; Zhimulev et al. 1981). These studies, which identify all of the genes in a small region by inducing and isolating mutations representing these genes, indicate that there are between 5000 and 6000 genes in D. melanogaster. Obviously, D. melanogaster has much more single copy DNA than is necessary to code for 6000 proteins, even including introns and other non-translated sequences. What then is the function of the excess DNA? How much is involved in

gene regulation and how are regulatory sequences positioned relative to the structural elements that they control? Are genes separated by large non-coding regions or are they clustered, resulting in "geneless" regions? Also, are genes which are co-regulated or which have similar functions located next to each other in the genome? The resolution of these questions will require extensive genetic and molecular analyses of several regions.

Saturation mapping of the region around rosy, the structural gene for xanthine dehydrogenase (Hilliker et al. 1981) was done in conjunction with intra-genic mapping experiments of rosy alleles to study the organization of the structural and regulatory elements of this gene (Chovnick et al. 1977; McCarron et al. 1979). This very elegant analysis has identified a regulatory or control region located between the XDH structural element and the gene adjacent to the left of rosy, 1(3)S12. The size of the XDH structural element is .005-map units which corresponds to approximately 4,100 bp of DNA while the control region, defined as the interval between 1(3)S12 and rosy, is .0016-map units or 1,310 bp of DNA (McCarron et al. 1979). The rosy gene has been cloned and a molecular analysis of this gene and the surrounding region is in progress (A. Chovnick, personal communication). Similar genetic and molecular analyses are underway in other regions of the D.melanogaster genome (Woodruff and Ashburner 1979; Benyajati et al. 1980; Judd et al. 1972; Bingham et al. 1981).

As promising as the experimental systems in D. melanogaster are for investigating genomic organization and gene expression, it is important that these types of analyses be done in several higher eucaryotes for comparison. In this thesis I examine the organization of the genetic material in the free-living nematode, Caenorhabditis elegans, by genetically characterizing a small region of the genome of this organism. C. elegans is an excellent candidate for this study since it is suitable for both genetic and molecular analyses. Its advantages as an experimental organism include a short generation time, a large number of progeny, a small genome size and a small number of cells.

C. elegans is a simple organism consisting of 811 and 971 somatic cells in the adult hermaphrodite and male, respectively (Sulston and Horvitz 1977). Although there are relatively few cells, these cells differentiate into nerve, hypodermis, muscle, gut and gonad. The complete cell lineage of this nematode, from zygote to adult, has been determined (Sulston and Horvitz 1977; von Ehrenstein and Schrierenburg 1979; J. Sulston, personal communication). The suitability of C. elegans for genetic analysis is described in Brenner (1974), Riddle (1978) and Herman and Horvitz (1981). The generation time of C. elegans is 3.5 days at 20°C. As indicated above, C. elegans has two sexes, hermaphrodite amd male. The chromosome constitution of the hermaphrodite is 5AA:XX while the males have 5 autosomes but only one sex chromosome (Brenner 1974). Hermaphrodites reproduce

by either self-fertilization or by mating with the males, and each hermaphrodite produces about 350 progeny. Over 430 genes have been identified and mapped to the six linkage groups. In addition, several chromosomal aberrations (deletions, duplications and translocations) have been isolated and characterized (Herman et al. 1976, Herman 1978; Herman et al. 1979; Meneely and Herman 1979; Moerman and Baillie 1981; Rose and Baillie 1980; Rosenbluth and Baillie 1981; Rogalski et al. 1982). The haploid genome of C. elegans contains approximately 8 x 10^7 base pairs of DNA: 17% repetitive and 83% single copy (Sulston and Brenner 1974). Several C. elegans genes have been cloned and characterized by restriction mapping and DNA sequencing (MacLeod et al. 1981; Kramer et al. 1982; Karn et al. 1983; T. Snutch, unpublished results).

One advantage C. elegans has over most other organisms is its suitability for studying muscle structure, function and regulation (Zengel and Epstein 1980a). Several genes affecting muscle structure have been identified (Waterston et al. 1980) including the genes coding for paramyosin (Waterston et al. 1977) and the major body-wall myosin heavy chain (MacLeod et al. 1977). The work described in this thesis is part of an ongoing analysis at both the genetic and molecular levels to study (1) the regulation of specific muscle genes during development in C. elegans and (2) the organization of the C. elegans genome around these genes (Moerman and Baillie 1979; Rose and Baillie 1980; Rogalski et al. 1982). Specifically, this thesis describes the

genetic analysis of a small region around unc-22 IV, a gene affecting muscle structure in C. elegans (Brenner 1974; Waterston et al. 1980; Zengel and Epstein 1980b; Moerman 1980).

Although the unc-22 gene product has not yet been identified, there is evidence indicating that this gene is involved in muscle structure or function. Mutations in unc-22 cause a characteristic twitching pattern along the body-wall musculature of the nematode. Circumstantial evidence indicates that this twitch is not due to abnormal neuronal input to the muscle, but rather is a result of a disorder in the muscle cell itself (Lewis et al. 1980; Moerman 1980). The body-wall musculature of unc-22 mutants appears disorganized when viewed under the electron microscope (Waterston et al. 1980; Moerman 1980). In addition, several unc-22 mutants are suppressed by a particular class of mutations in unc-54 I, the gene that codes for the major body-wall myosin heavy chain in C. elegans (Moerman 1980; Moerman et al. 1982). This suppression by unc-54 mutations indicates that the unc-22 gene product interacts with this major muscle protein.

The first objective of this thesis was to identify essential genes in a defined region around unc-22. This was accomplished by isolating lethal mutations linked to unc-22, positioning these mutations by recombination and deficiency mapping and identifying lethal complementation groups. Several nonessential genes in this region have also been positioned. The eventual goal of this analysis is to determine the total number

of genes in this small, defined interval of Lnkage Group IV (LG IV). The only other small genomic region in C. elegans that has been characterized in this manner is the unc-15 I region (Rose and Baillie 1980).

A second objective of the thesis was to refine and expand the existing fine-structure map of the unc-22 gene (Moerman and Baillie 1979; Moerman 1980) in order to define its structural and regulatory elements. To this end, intra-genic mapping experiments were performed to position, more accurately, previously mapped alleles and to position new alleles within the unc-22 fine-structure map.

B. MATERIALS AND METHODS

The genetic nomenclature follows the recommendations of Horvitz et al. (1979).

I. Nematode strains and culture conditions

The wild-type (N2) strain and strains carrying dpy-18(e364)III, unc-5(e152)IV, unc-43(e266)IV, unc-31(e169)IV, unc-30(e318)IV, dpy-4(e1166)IV, unc-46(e177)V and dpy-11(e224)V were originally obtained from the Caenorhabditis elegans (Maupas, 1900) Dougherty, 1953, var. Bristol, stock collection at the Medical Research Council, Cambridge, England. The dpy-20(e1282)IV and unc-5(e152)IV/nT1; dpy-11(e224)V/nT1 strains were obtained from the Caenorhabditis Genetics Center at the University of Missouri, Columbia. The following mutations and deficiencies of the unc-22 gene were isolated at Simon Fraser University: s7, s8, s12, s18, and s32 (Moerman and Baillie 1979), s35, s36, sDf2, sDf7, sDf8, sDf9, sDf10 and sDf19 (Moerman and Baillie 1981), and s699 and s700 (R. Rosenbluth, unpublished results). All of the unc-22 mutations and deficiencies have a recessive Unc-22 (twitcher) phenotype except sDf19, which has a dominant twitcher phenotype. The deficiencies, including sDf19, are recessive lethals. A genetic map of C. elegans is shown in Appendix 1.

The lethal mutations and deficiencies were maintained in heterozygous strains of the genotypes let-a(sx)unc-22(s7)/++ and sDfy/+, respectively. In the case of sDf19, the heterozygous strain was maintained by selecting Unc-22 (twitcher) progeny each generation. I was able to maintain the other strains without losing the lethal or deficiency carrying chromosome because worms heterozygous for a mutation or deficiency in the unc-22 gene can be distinguished from homozygous wild-type worms in a 1% nicotine solution (nicotine alkaloid, Sigma). The heterozygotes twitch in the nicotine solution while the homozygous wild-types do not (Moerman and Baillie 1979, 1981). Nematodes were cultured on Nematode Growth Media (NGM) streaked with E. coli OP50 as described by Brenner (1974).

II. Isolation of unc-22-linked lethal and sterile mutations

Twenty-six of the unc-22-linked lethal and sterile mutations that were used in this study were isolated by D. Moerman as described in Moerman (1980) and Rogalski et al. (1982). The remaining 26 mutations were isolated in this study using a similar protocol. Homozygous unc-22(s7) hermaphrodites were treated with 0.025M ethyl methanesulfonate (EMS) (Brenner 1974) and then mated to wild-type males. Young, adult F1 heterozygous hermaphrodites were placed on 35mm Petri plates (one per plate) at 20°C for 24 hr and then removed. The F2 progeny on each plate was screened for fertile adult Unc-22

(twitcher) worms. Fertility was established by observing internal eggs. C. elegans is transparent, therefore, fertilized eggs can be observed inside the hermaphrodite under a dissecting microscope. A lethal mutation on the unc-22(s7) marked chromosome was identified by the absence of adult twitchers and a putative lethal strain was established by selecting F2 heterozygotes from the same plate (using nicotine: see above). A sterile mutation was identified by the presence of sterile adult twitchers and these mutations were treated as "lethals". The putative lethal strains were retested at 20° C and confirmed lethals were maintained at 15° C.

III. Re-isolation of gamma-ray induced unc-22 mutations

The s699 and s700 alleles of unc-22 were induced with 1500 Rads gamma radiation in a dpy-18(e364)III/eT1; unc-46(e177)V/eT1 strain (R. Rosenbluth, personal communication). The s699 mutation was re-isolated from the dpy-18(e364), unc-46(e177) and eT1(III;V) mutations by the following procedure. 1) The s699IV; dpy-18III/eT1; unc-46V/eT1 hermaphrodites were mated to dpy-18/+ males and Dpy progeny were picked. 2) These Dpy hermaphrodites (dpy-18; s699/+; unc-46/+) were allowed to self cross and Dpy Unc-22 progeny were picked. 3) The Dpy Unc-22 hermaphrodites (dpy-18; s699; unc-46/+ or +/+) were then mated to wild-type males. The F1 heterozygotes were picked and their progeny were screened for Unc-22 individuals. 4) The Unc-22 hermaphrodites

(dpy-18/+ or +/+; s699; unc-46/+ or +/+) were allowed to self cross and a strain was established from an individual which segregated only Unc-22 progeny. 5) To ensure that the new strain did not carry the unc-46 mutation, hermaphrodites were mated to unc-46/+ males. No Unc-46 outcross progeny were observed.

The s700 mutation was re-isolated in the same manner except that when the original strain was mated to dpy-18/+ males no F1 Dpy progeny were observed. Instead, the F2 progeny were screened for Dpy individuals which were picked and then treated in the same manner as the F1 Dpy progeny described above.

IV. Construction of strains

The double mutant strains used in this thesis have the following genotypes: unc-5(e152)dpy-4(e1166)IV, unc-5(e152)unc-22(s18)IV, dpy-20(e1282)unc-22(s7)IV, unc-43(e266)unc-22(s18)IV, unc-22(s7)unc-31(e169)IV, and unc-22(s12)unc-30(e318)IV. The unc-43unc-22 strain was constructed for this study in the following manner. Hermaphrodites homozygous for unc-43 were mated to unc-22/+ males. F1 hermaphrodites with the genotype unc-43 +/+ unc-22 were selected using nicotine and their progeny were screened for Unc-43 hermaphrodites that twitched in nicotine (unc-43 + /unc-43unc-22). These individuals segregated homozygous unc-43unc-22 progeny which were picked to establish a strain. The unc-5unc-22 and dpy-20unc-22 strains were also constructed

for this study using the same procedure while the three remaining strains were obtained from the laboratory strain collection and had been constructed by Moerman (1980).

Three strains of the genotype unc-5(e152)unc-22(sx)dpy-4(e1166)IV , carrying the s8, s12 and s18 alleles of the unc-22 gene, were used in the fine-structure mapping experiments. These strains were obtained from D. Moerman and their construction is described in Moerman and Baillie (1979).

An sDf19IV/+; dpy-11(e224)V strain was constructed by mating sDf19/+ hermaphrodites with dpy-11/+ males and selecting F1 Unc-22 progeny. The Unc-22 progeny from this cross will be either sDf19/+; dpy-11/+ or sDf19/+; +/+ because of the dominant phenotype of sDf19 (see above). Dpy-11 Unc-22 hermaphrodites were selected from the F2 generation and a strain was maintained by selecting Dpy Unc-22 hermaphrodites every generation.

V. Positioning mutations by recombination mapping

Lethal mutations were positioned relative to unc-22 by the following protocol. The progeny of unc-5(e152) + + dpy-4(e1166)/+ unc-22(s7)let-a(sx) + hermaphrodites were screened for Unc-22 recombinants (either unc-5unc-22/+ unc-22 or unc-22dpy-4/unc-22 +) and these were progeny tested. If the Unc-22 recombinants segregated Unc-22 Dpy-4 progeny then the lethal mutation must be to the right of unc-22, whereas if the

Unc-22 recombinants segregated Unc-5 Unc-22 progeny the lethal mutation must be to the left. Nonessential genes in the region were positioned relative to the unc-22 gene by 3-factor mapping. The progeny of unc-x unc-22 +/ + + dpy-4 hermaphrodites were screened for both Unc-22 and Unc-x recombinants which were progeny tested. If the Unc-22 recombinants segregated Dpy-4 Unc-22 progeny then unc-x must be to the right of unc-22. Conversely, if the Unc-x recombinants segregated Dpy Unc-x progeny then unc-x must be to the left of unc-22.

The map distance from unc-22 to a lethal site was determined by measuring the recombination frequency between unc-22 and the lethal. Mapping was done at 20°C and all of the progeny were scored for recombinant events as described by Rose and Baillie (1979). Ten, 20 or 30 late larval hermaphrodites (let-a(sx)unc-22(s7)/++) were placed on small(35mm) or medium(60mm) Petri plates, one per plate, and transferred every 6-18 hr. The F1 phenotypes were scored and the recombination frequency was calculated using the formula $p = 1 - \sqrt{1 - [3U/(U + W)]}$ where p = recombination frequency, U = number of Unc-22 progeny and W = number of phenotypically wild-type progeny. The derivation of this mapping equation is described in Appendix 2.

In a similar manner, the recombination frequency from unc-22 to several of the nonessential genes in the region was determined by placing 10 or 20 unc-22unc-x/++ late larval hermaphrodites on plates, one per plate, and transferring them every 6-18 hr. The F1 progeny were scored and the recombination

frequency was calculated.

VI. Positioning mutations by deficiency mapping

The lethal-bearing strains were complemented with the deficiencies by mating males heterozygous for a lethal mutation, let-a(sx)unc-22(s7)/++, to hermaphrodites carrying the deficiency, sDfy/+. The progeny of this cross were screened for males, to confirm that the hermaphrodites had outcrossed, and for the presence of adult male and hermaphrodite twitchers. When present, several hermaphrodite twitchers were put on a separate plate and examined for fertility. If no fertile, adult hermaphrodite twitchers were found, the let locus was considered to be included in the deficiency.

Complementation tests between visible mutations and deficiencies were done by crossing sDfx/+ hermaphrodites to the heterozygous mutant males. The progeny were screened for males with the mutant phenotype. The presence of these mutant males indicated that the deficiency uncovered the visible site.

Complementation tests between the lethal mutations and the sDf19 deficiency were performed using a different procedure because of the dominant twitcher phenotype of the sDf19 mutation. Hermaphrodites with the genotype, sDf19IV/+; dpy-11(e224)V, were mated to let-a(sx)unc-22(s7)/++ males and the outcrossed Unc-22 hermaphrodites (either sDf19/+; dpy-11/+ or sDf19/let-aunc-22; dpy-11/+) were picked and progeny tested.

If none of the F1 hermaphrodites from a particular cross carried the let-aunc-22 chromosome then the lethal mutation was considered to be included in the deficiency.

VII. Identification of lethal complementation groups

The procedure for the complementation tests was modified from that used by Brenner (1974) for visible mutations. Males of a lethal strain (let-a(sx)unc-22(s7)/++) were crossed with hermaphrodites of another lethal strain (unc-22(s7)let-b(sy)/++). The progeny of this cross were screened for males, to confirm that the hermaphrodites had outcrossed, and for the presence of adult male and hermaphrodite twitchers. When present, several hermaphrodite twitchers were put on a separate plate and examined for fertility. An absence of fertile adult hermaphrodite twitcher progeny indicated that the two lethals failed to complement and were alleles of the same gene.

VIII. Characterization of unc-22-linked lethal mutations

The lethal mutations were characterized by determining their effective lethal phases (i.e. the stage at which a mutation blocks the development of the let-a(sx)unc-22(s7) homozygote). The effective lethal phases of the s41 to s47, s49 to s52, s59, s63, s65, s166, s168 to s176, s212 to s216 and s254

mutations were determined by D. Moerman as described in Moerman (1980) and Rogalski et al. (1982). The remaining 27 lethal mutations were characterized using essentially the same procedure. Young twitcher larvae were removed from the lethal stock plate, placed on another plate for 1 to 3 days and then measured. The larvae were heat killed (by placing the tip of a hot soldering iron in the agar beside them) and then measured with an ocular micrometer at 50X magnification. The final lengths of these doubly mutant worms was used to determine if they were embryonic, early, mid or late larval lethals. The larval lethals were not assigned to a specific larval developmental stage because no record of their molts during development was made.

The deficiencies sDf2, sDf7, sDf8, sDf9 and sDf10 were characterized in a similar manner. Several sDfx/+ hermaphrodites were placed on a plate, allowed to lay eggs for approximately 2 hr and then removed. The plates were screened after 24 and 48 hr for unhatched eggs or twitcher larvae.

IX. Mapping alleles within the unc-22 gene

The procedure used for the fine-structure mapping of the unc-22 gene was modified from that of Moerman and Baillie (1979). Thirty to 100 young adult hermaphrodites of the genotype unc-5(e152)unc-22(sx)dpy-4(e1166)/+ unc-22(sy) + were placed on large (100mm) petri plates and transferred every 6 to 12 hr

until there were approximately 200 plates. The plates were incubated at 15 °C to 20° C until the F2 generation was mature (6 days at 20° C) and then were screened for individuals that did not twitch (either wild-type, Dpy, Dpy Unc-5 or Unc-5). All nonUnc-22 recombinants were picked and confirmed as such by progeny testing. Since the hermaphrodites used in the fine-structure mapping experiments were heterozygous for the flanking markers, unc-5 and dpy-4, the relative positions of the two unc-22 alleles being mapped could be determined. If sy is to the left of sx then the recombinant unc-22+ chromosome will carry the unc-5 marker and recombinant individuals will be phenotypically wild-type (unc-5 + +/+ sy +) or Unc-5 (unc-5 + +/unc-5 sx dpy-4). Alternatively, if sy is to the right of sx then the dpy-4 marker will be carried on the recombinant chromosome and the nonUnc-22 recombinants will be either wild-type (+ + dpy-4/ + sy +) or Dpy-4 (+ + dpy-4/unc-5 sx dpy-4).

The map distance between two unc-22 alleles was calculated using the formula d = 2(W)x100/T from Moerman and Baillie (1979) where d = map distance; W = nonUnc-22 recombinants and T = total F2 progeny. The total number of F2 progeny was estimated in the following manner. First, the total number of F1 individuals was estimated by counting all of the F1 progeny on a sample (20%) of the plates. This number was then halved to give the number of F1 heterozygotes, since the recombination event can only occur in these individuals. Next, the total number of progeny produced by

an unc-5unc-22(sx)dpy-4/+ unc-22(sy) + heterozygote was determined by completely scoring all of the progeny of 10 such hermaphrodites. However, not all of the progeny of an F1 heterozygote are screened in these experiments, since screening occurs at low magnifications (6X and 12X) and only recombinant individuals which were at the L3 or later stages of development would be observed. Approximately one-half of the F2 progeny would be at the appropriate developmental stages at the time the plates were screened. Therefore, the total number of F2 progeny that were screened was estimated as the number of F1 heterozygotes multiplied by one-half the total progeny of a heterozygous hermaphrodite.

The sDf19/unc-5(e152)unc-22(sx)dpy-4(e1166) strains used for intra-genic mapping of the dominant unc-22 mutation, sDf19, were constructed as follows. Males of the genotype unc-5unc-22dpy-4/+++ were mated to sDf19/+ hermaphrodites and F1 twitchers (either sDf19/+++ or sDf19/unc-5unc-22dpy-4) were selected and their genotype confirmed by progeny testing. Strains were established and maintained for several generations by picking Unc-22 hermaphrodites from plates containing unc-5unc-22dpy-4 homozygotes.

The procedure for mapping the breakpoint of sDf19 inside the unc-22 gene was identical to that described above. Fifty unc-5unc-22(sx)dpy-4/+ sDf19 + hermaphrodites were placed on large plates, transferred at 12 hr intervals until there were 200 plates and the F2 generation was screened for nonUnc-22

recombinants. The map distance from the deficiency breakpoint to the unc-22 allele was calculated using the formula described above except that the number of nonUnc-22 recombinants was multiplied by 4 since half of the recombinants would not be observed due to the dominant twitcher phenotype of the sDf19/+ heterozygote.

X. Construction of a balanced lethal strain

A let-71(s692)unc-22(s7)IV/nT1 strain was constructed using the following protocol. Lethal-bearing males, let-71unc-22/++, were mated to hermaphrodites carrying the translocation, unc-5(e152)IV/nT1; dpy-11(e224)V/nT1. F1 hermaphrodites that twitched in nicotine (either let-71unc-22/nT1; +/nT1 or + let-71unc-22/unc-5 + +; dpy-11/+) were picked and their genotype was confirmed by progeny testing. A let-71unc-22/nT1; +/nT1 strain was established by selecting wild-type hermaphrodites from plates which did not contain Dpy, Unc-5 or Dpy Unc-5 worms. This strain was maintained by selecting wild-type hermaphrodites.

XI. Construction of strains for DNA hybridization

Several strains were constructed during this study that were used by other researchers to position cloned C. elegans DNA fragments on LG IV (Rose et al. 1982; T.Snutch, personal

communication; D.Baillie and K.Beckenbach, personal communication). The first strain was constructed as follows. Heterozygous unc-22(s18)/+ males of the Bristol (N2) variety of C. elegans were crossed to Bergerac (BO) variety hermaphrodites and F1 hermaphrodites of the genotype unc-22/+ were selected using nicotine. The Unc-22 progeny produced by these heterozygotes will be homozygous for the N2 fourth chromosome (N2 IV) but will vary as to whether they are homozygous N2, heterozygous N2/BO or homozygous BO with respect to the other five chromosomes. Four hundred F2 Unc-22 progeny were selected to establish a strain that was homozygous for N2 IV and heterozygous N2/BO for the other chromosomes.

The second strain, which was homozygous for the BO fourth chromosome (BO IV), was established at the same time. In this case, the homozygous BO IV hermaphrodites were selected in the F2 generation as worms that did not twitch in nicotine.

The remaining strains were constructed by mating unc-22(s7)unc-31(e169)/++ N2 males to BO hermaphrodites and selecting F1 individuals with the genotype unc-22unc-31/++. The F2 generation was screened for Unc-31 recombinants (+ unc-31/unc-22unc-31) and the homozygous + unc-31 progeny of these hermaphrodites were identified using nicotine. Homozygous + unc-31 strains were established for each of the six recombinants that were selected. The + unc-31 recombinant chromosomes carried in these strains have the Bergerac fourth chromosome to the left of the crossover point and the Bristol

fourth chromosome to the right of this point.

C. RESULTS

I. Isolation of unc-22-linked lethal and sterile mutations

In addition to the lethal and sterile mutations which had been previously isolated, another set of mutations was isolated in this study. Of 1544 F1 individuals whose progeny were examined, 36 were found to carry a lethal or sterile mutation linked to unc-22(s7). Ten of these original 36 mutant strains were discarded because of either poor fecundity or loose linkage to unc-22. Lethal mutations that were loosely linked to the unc-22 locus could not be effectively maintained in this system. The induction rate of lethal and sterile mutations in the region defined by sDf2 was approximately 9.7×10^{-3}. These results are summarized in Table 1.

Table 1

Results of screen to isolate *unc-22*-linked lethal and sterile mutations in *Caenorhabditis elegans*

Dose of EMS used	Number of F1's screened	Number of lethal mutations inside *sDf2*	outside *sDf2*	discarded	Mutation rate in *sDf2* region
0.025M	1544	15	11	10	9.7×10^{-3}

II. Initial positioning of four deficiencies

A preliminary genetic analysis of the unc-22 region of linkage group IV (Moerman 1980) had identified and positioned eight essential genes: let-51 to let-56, let-59 and let-60 and two deficiencies, sDf1 and sDf2, in the region around the unc-22 gene. These eight let loci, as well as unc-30 and unc-31 were used to position four additional deficiencies (sDf7, sDf8, sDf9 and sDf10) that delete the unc-22 locus. These latter deficiencies had been generated with formaldehyde in a previous study (Moerman and Baillie, 1981). The results of the complementation tests between the deficiencies and the mapped loci are given in Table 2. The positions of the deficiencies are shown in Figure 1. The results mean that the sDf2 region can be divided into several smaller regions by the breakpoints of these four deficiencies.

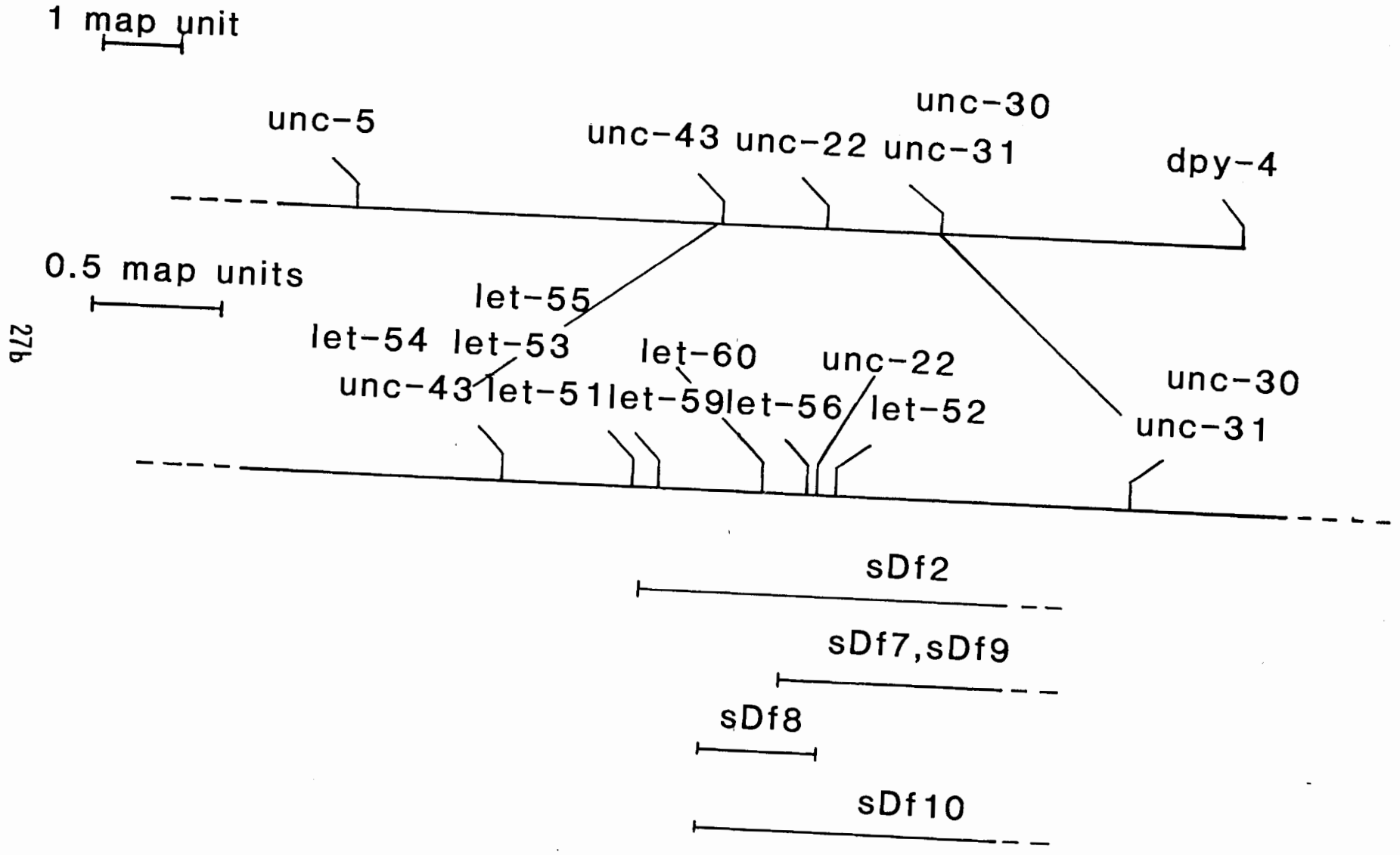
Table 2

Complementation tests between eight essential genes and four deficiencies in the unc-22 IV region of Caenorhabditis elegans*

Genes	Alleles	Deficiencies			
		sDf7	sDf8	sDf9	sDf10
let-54	s44	+	+	+	+
let-53	s43	+	+	+	+
let-55	s45	+	+	+	+
let-51	s41	+	+	+	+
let-59	s49	+	+	+	+
let-60	s59	+	-	+	-
let-56	s46	-	-	-	-
let-52	s42	-	+	-	-

*Complementation is indicated by + and noncomplementation is indicated by -.

Figure 1. The region around unc-22 IV in Caenorhabditis elegans. The relative positions of nonessential genes in this region are shown on the top map. The bottom map is an expansion of the region between unc-43 and unc-31 showing the positions of eight essential genes relative to unc-22 and the positions of five deficiencies.


1 map unit
unc-5
unc-43
unc-22
unc-31
unc-30
dpy-4
0.5 map units
let-55
let-54
let-53
let-60
unc-22
unc-43
let-51
let-59
let-56
let-52
unc-30
unc-31
sDf2
sDf7,sDf9
sDf8
sDf10

27b

III. Positioning mutations by recombination mapping

The first step in mapping the lethal and sterile mutations was to determine the position of each mutation relative to unc-22 IV. After a lethal was positioned, the map distance from the lethal site to the unc-22 gene was then determined. Fifty-three mutations were mapped in this manner, including four (s42, s46, s49 and s59) of the eight mutations previously positioned by Moerman (1980) and one mutation, s504, which was isolated and positioned by D. Pilgrim (personal communication). The results of the recombination mapping are summarized in Tables 3 and 4. Forty-four of the lethal mutations mapped to the left of unc-22 whereas only eight mapped to its right. The position obtained for s59 in this thesis differs from that obtained by Moerman (1980). let-60(s59) was initially positioned 0.5-map units to the right of unc-22. However, the results in Table 3 place s59 0.2-map units to the left of unc-22. Several nonessential loci in the region were also positioned relative to unc-22 (Table 5; Figure 1).

Table 3

Left-right positioning and Two-factor mapping results for unc-22-linked lethal and sterile mutations uncovered by the deficiency, sDf2 IV, in Caenorhabditis elegans

Mutation	Position relative to unc-22	Phenotypes Unc-22	Phenotypes Wild	Map units from unc-22*
s42	right	3	6827	0.07 (0.01-0.19)
s46	left	1	5806	0.03 (0.001-0.14)
s49	left	8	2127	0.56 (0.24-1.10)
s50	left	0	2342	-
s52	left	9	1423	0.94 (0.43-1.79)
s59	left	3	2097	0.21 (0.04-0.62)
s65	left	13	2108	0.92 (0.48-1.58)
s168	left	1	1919	0.08 (0.02-0.43)
s170	left	8	1855	0.65 (0.28-1.27)
s171	left	6	1360	0.66 (0.24-1.44)
s172	left	12	2257	0.80 (0.41-1.39)
s173	left	1	5041	0.03 (0.001-0.17)
s174	left	5	1793	0.42 (0.14-0.98)
s175	left	12	2394	0.75 (0.38-1.30)
s176	right	3	2008	0.22 (0.04-0.63)
s214	right	7	4601	0.23 (0.09-0.47)
s216	left	9	2259	0.60 (0.27-1.14)
s254	left	3	1766	0.25 (0.05-0.74)

Table 3 cont'd

Mutation	Position relative to unc-22	Phenotypes Unc-22	Phenotypes Wild	Map units from unc-22*
s504	left	4	5632	0.11 (0.03-0.27)
s677	left	4	1980	0.30 (0.08-0.77)
s678	left	7	2279	0.46 (0.18-0.94)
s679	left	9	2112	0.64 (0.29-1.21)
s680	right	8	2250	0.53 (0.23-1.05)
s681	left	8	2360	0.51 (0.22-0.99)
s684	left	16	2011	1.19 (0.61-1.77)
s685	left	9	2148	0.63 (0.29-1.19)
s689	left	18	2521	1.07 (0.58-1.56)
s692	left	19	2478	1.15 (0.63-1.66)
s693	right	8	2301	0.52 (0.22-1.03)
s694	left	5	1785	0.42 (0.14-0.97)
s695	left	14	2261	0.92 (0.51-1.55)
s696	right	3	1524	0.29 (0.06-0.86)
s697	left	6	1542	0.58 (0.21-1.27)
s698	right	1	4185	0.04 (0.001-0.16)

*Map units = 100p and $p = 1 - \sqrt{1 - [3U/(U+W)]}$ where p = recombination frequency, U = number of Unc-22 progeny and W = number of Wild progeny.

95% confidence intervals are shown in the brackets and were calculated using the equation $1.96\sqrt{Npq}$ or using the Stevens table.

Table 4

Left-right positioning and two-factor mapping results for *unc-22*-linked lethal mutations outside the region uncovered by the deficiency, *sDf2 IV*, in *Caenorhabditis elegans*

Mutation	Position relative to *unc-22*	Phenotypes Unc-22	Phenotypes Wild	Map units from *unc-22**
s47	left	12	2660	0.67 (0.35-1.18)
s51	left	21	1712	1.84 (1.07-2.60)
s63	left	41	1552	3.94 (2.76-5.12)
s166	left	29	1871	2.32 (1.48-3.15)
s169	left	26	1935	2.01 (1.25-2.77)
s212	left	2	1765	0.17 (0.02-0.61)
s213	left	18	4749	0.57 (0.31-0.83)
s215	left	6	1698	0.53 (0.19-1.16)
s673	left	22	2308	1.43 (0.83-2.02)
s674	left	39	2426	2.40 (1.65-3.15)
s675	left	14	1853	1.13 (0.62-1.90)
s676	left	33	2143	2.30 (1.52-3.08)
s682	left	55	1680	4.87 (3.62-6.13)
s683	left	22	1897	1.73 (1.02-2.45)
s686	left	13	2464	0.79 (0.42-1.35)

Table 4 cont'd

Mutation	Position relative to *unc-22*	Phenotypes Unc-22	Wild	Map units from *unc-22**
s687	left	37	1724	3.20 (2.18-4.22)
s688	left	9	1660	0.81 (0.37-1.54)
s690	right	1	1928	0.08 (0.002-0.43)
s691	left	33	1977	2.49 (1.65-3.34)

*Map units = 100p and $p = 1 - \sqrt{1 - [3U/(U+W)]}$ where p = recombination frequency, U = number of Unc-22 progeny and W = number of Wild progeny.

95% confidence intervals are shown in the brackets and were calculated using the equation $1.96\sqrt{Npq}$ or using the Stevens table.

Table 5

Two-factor mapping results for nonessential genes in the unc-22 IV region in Caenorhabditis elegans

Genotype of P_0 a b/ + +	F1 Phenotypes a	b	Wild	Map distance between a and b*
unc-5(e152)dpy-4(e1166)/ + +	196	167	2337	11.46 (10.40-12.54)**
unc-5(e152)unc-22(s18)/ + +	52	29	2106	2.87 (2.26-3.48)
unc-43(e266)unc-22(s18)/ + +	28	35	4225	1.12 (0.84-1.39)
dpy-20(e1282)unc-22(s7)/ + +	5	7	4186	0.22 (0.11-0.36)
unc-22(s7)unc-31(e169)/ + +	17	26	2926	1.10 (0.77-1.43)
unc-22(s12)unc-30(e318)/ + +	17	31	2777	1.29 (0.93-1.65)
unc-31(e169)unc-26(e345)/ + +	27	-	3489	1.16 (0.73-1.60)***
unc-31(e169)dpy-4(e1166)/ + +	53	74	1976	4.78 (3.98-5.58)***

*Map distance = 100p where $p = 1 - \sqrt{1 - \left[3(a+b)/\ 2Wild + (a+b)\right]}$.

**95% confidence intervals are shown in the brackets and were calculated using the equation $1.96\sqrt{pqN}$ or using the Stevens table.

***These map distances were determined by D. Baillie and G. Wild.

IV. Positioning mutations by deficiency mapping

The 53 lethal and sterile mutations that were positioned by recombination mapping were tested for complementation with sDf2. Nineteen of these, 18 of which map to the left of unc-22, complemented sDf2 and, therefore, lie outside its region. The number of essential genes defined by these 19 mutations was not determined nor was any further mapping done with these lethals.

The remaining 34 lethal and sterile mutations failed to complement sDf2. The results of complementation tests with the deficiencies, sDf7, sDf8, sDf9 and sDf10, localized these 34 mutations into six separate zones. Table 6 summarizes the results of these tests. Inter se complementation tests were done between lethals that occupied the same zone, and the 20 genes defined by these tests are shown in the first column of Table 6. Figure 2 shows the positions of these genes. Sixteen lethals are in zone 1, and they define 12 loci, let-69(s684), let-70(s689) let-71(s692), let-61(s65), let-72(s50,s695), let-59(s49,s172), let-62(s175), let-63(s170,s679), let-64(s171,s216), let-73(s685), let-74(s697) and let-91(s678). There are four lethals in zone 2 that define two genes, let-65(s174,s254,s694) and let-60(s59). This zone also contains the dpy-20 gene. The two lethal mutations in zone 3, s504 and s677, are alleles of let-92. Similarly, the four mutations in zone 4, s46, s52, s168 and s173, are all alleles of let-56. The unc-22 locus is also in

this zone. Zone 5 contains s42, which defines the let-52 locus, whereas zone 6 has five mutations that define three loci, let-66(s176) , let-67(s214) and let-68(s680,s693,s696). Lethal complementation groups located in the same zone were not positioned with respect to each other. Sixteen of the let loci map to the left of unc-22, and four map to the right of it. The lethal, sterile and visible mutations so far uncovered by sDf2 represent 22 loci in a 1.5-map unit interval.

V. Positioning the deficiencies

As well as positioning the lethal mutations, the complementation tests between the deficiencies and the lethals also refined the positions of the deficiency breakpoints. The results in Table 6 show that sDf7, sDf8, sDf9 and sDf10 are all smaller than sDf2 and have both of their breakpoints in the region uncovered by sDf2. The two smallest deficiencies, sDf7 and sDf8, are not more than 0.3- and 0.4-map units in length, respectively whereas sDf8 extends approximately 0.4- to 0.6-map units and sDf10 extends approximately 0.4- to 0.8 map-units. There are at present three genes, let-56, unc-22 and let-52, in the region uncovered by the smallest deficiency.

Table 6

Complementation tests between genes and deficiencies in the unc-22 IV region of Caenorhabditis elegans*

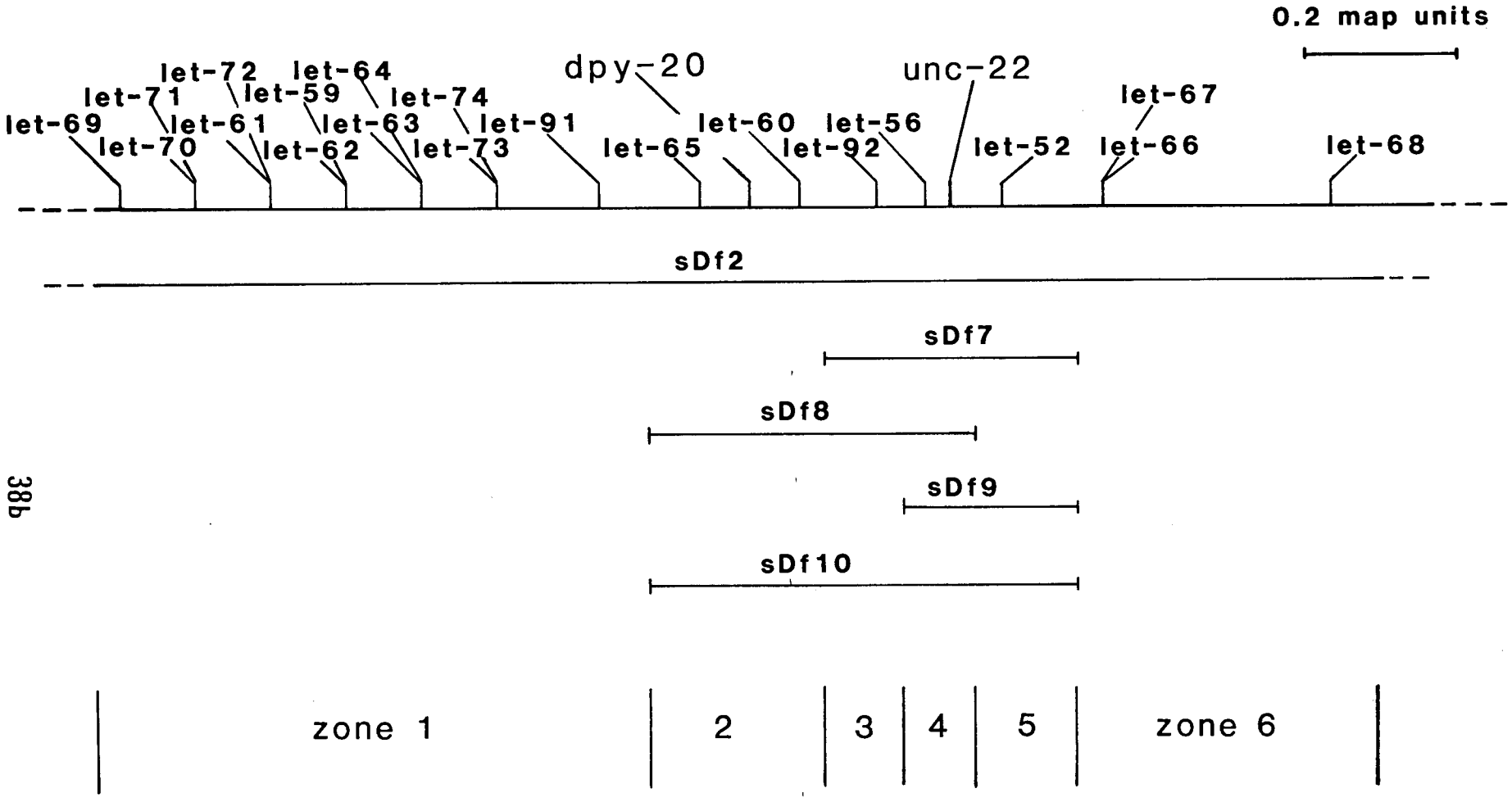
		Deficiencies				
Genes	Alleles	sDf7	sDf8	sDf9	sDf10	Zone
let-69	s684	+	+	+	+	1
let-70	s689	+	+	+	+	1
let-71	s692	+	+	+	+	1
let-61	s65	+	+	+	+	1
let-72	s52, s695	+	+	+	+	1
let-59	s49, s172	+	+	+	+	1
let-62	s175	+	+	+	+	1
let-63	s170, s679	+	+	+	+	1
let-64	s171, s216	+	+	+	+	1
let-73	s685	+	+	+	+	1
let-74	s697	+	+	+	+	1
let-91	s678	+	+	+	+	1
let-65	s174, s254, s694	+	-	+	-	2
let-60	s59	+	-	+	-	2
dpy-20	e1282	+	-	+	-	2

Table 6 cont'd

Genes	Alleles	Deficiencies				Zone
		sDf7	sDf8	sDf9	sDf10	
let-92	s504, s677	-	-	+	-	3
let-56	s46, s50, s168, s173	-	-	-	-	4
unc-22	s7	-	-	-	-	4
let-52	s42	-	+	-	-	5
let-66	s176	+	+	+	+	6
let-67	s214	+	+	+	+	6
let-68	s680, s693, s696	+	+	+	+	6

*Complementation is indicated by + and noncomplementation is indicated by -.

Figure 2. Essential genes, nonessential genes and deficiencies in the region of linkage group IV in Caenorhabditis elegans uncovered by sDf2. The positions of the deficiencies are shown below the line. The six zones defined by the deficiency endpoints are also shown.


0.2 map units
let-72
let-64
dpy-20
unc-22
let-71
let-59
let-74
let-67
let-69
let-61
let-63
let-91
let-60
let-56
let-70
let-62
let-73
let-65
let-92
let-52
let-66
let-68
sDf2
sDf7
sDf8
sDf9
sDf10
zone 1
2
3
4
5
zone 6

VI. Effective lethal phases

The various lethal homozygotes that were isolated and characterized in this study exhibited a wide range of terminal phenotypes, from the embryonic stage to the sterile adult (Tables 7 and 8; Figure 3). sDf2 homozygotes are embryonic lethals, but the remaining four deficiency strains develop to the first larval stage when homozygous. At present no clustering of genes affecting a particular developmental stage is apparent in this region. The sterile category accounts for 19% (5 of 26) of the lethals characterized in this study and 15% (4 of 26) of the lethals characterized by Moerman (1980).

Table 7

Characterization of lethal and sterile mutations uncovered by the deficiency, sDf2, in Caenorhabditis elegans

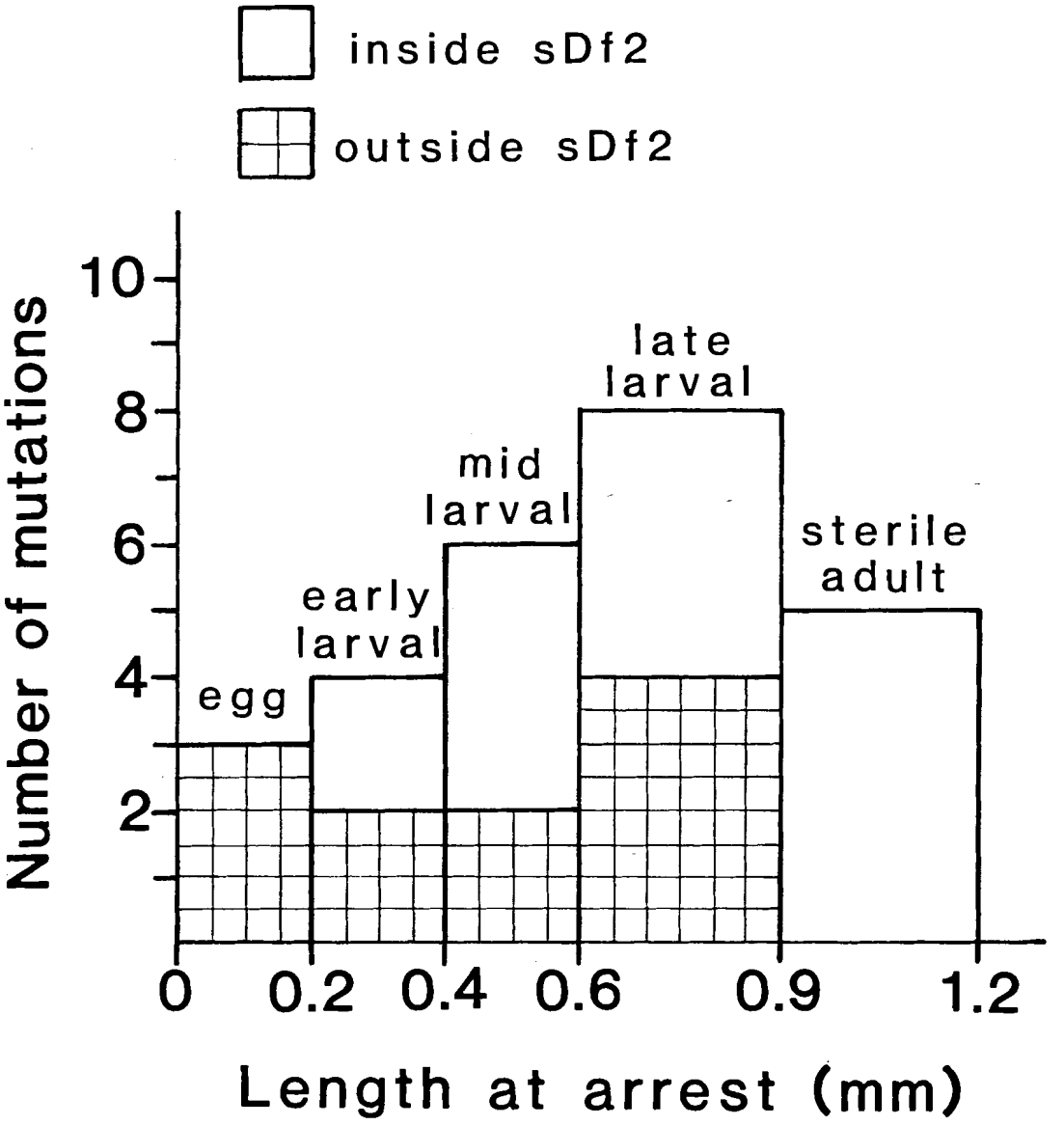
Genotype (all lets cis with unc-22(s7))	Hermaphrodite length at arrest (mm)	Effective lethal phase
s504	0.3	early larval
s677	0.3	early larval
s678	0.4	mid-larval
s679	0.7	late larval
s680	0.9	sterile
s681	0.3	early larval
s684	0.6	late larval
s685	1.0	sterile
s689	0.5	mid-larval
s692	0.9	sterile
s693	0.8	late larval
s694	0.9	sterile
s695	0.4	mid-larval
s696	1.1	sterile
s697	0.8	late larval (leaky)
s698	0.4	mid-larval

Table 8

Characterization of lethal mutations that lie outside the region uncovered by the deficiency, *sDf2*, in *Caenorhabditis elegans*

Genotype (all *lets* cis with *unc-22(s7)*)	Hermaphrodite length at arrest (mm)	Effective lethal phase
s673	0.3	early larval
s674	0.8	late larval
s675	0.6	late larval
s676	0.3	early larval
s682	-	egg
s683	-	egg
s686	0.6	mid-larval
s687	0.8	late larval
s688	0.5	mid-larval
s690	-	egg
s691	0.7	late larval

Figure 3. Distribution of the Effective Lethal Phases of unc-22 IV-linked mutations in Caenorhabditis elegans. The final lengths of the let-a(sx)unc-22(s7) hermaphrodites are shown for the twenty-six lethal and sterile mutations that were isolated in this study.


inside sDf2
outside sDf2
Number of mutations
10
8
6
4
2
egg
early larval
mid larval
late larval
sterile adult
0
0.2
0.4
0.6
0.9
1.2
Length at arrest (mm)

VII. Fine-structure mapping

Three intra-genic mapping experiments were performed in order to refine the positions of 3 previously mapped unc-22 alleles. The results of these mapping experiments are summarized in Tables 9 and 10 and in Figure 4.

The EMS-induced unc-22 allele, s32, is suppressed by the informational suppressors sup-5 III (Waterston and Brenner 1978) and sup-7 X (Waterston 1981) and is, therefore, in the structural element of the unc-22 gene (Moerman 1980; D. Moerman, personal communication). Prior to this analysis, s32 had been positioned only relative to s8, which is near the left boundary of the fine-structure map (Moerman 1980). This study has positioned s32 relative to s12, the right most allele of the unc-22 gene. Since no recombination occurred between these two alleles (Table 9) the site of the s32 mutation must be very close to the s12 site (Figure 4).

Four formaldehyde-induced unc-22 mutations had previously been positioned within the unc-22 fine-structure map between s8 and s12 which represent, respectively, the left and right boundaries of the map (Moerman 1980). I have mapped two of these mutations, s35 and s36, against s18, another allele of this gene. In both cases, no recombination occurred between the two alleles being tested (Table 9). These results suggest that both s35 and s36 map very close to s18. The new positions of these

two mutations are shown in Figure 4.

Two new gamma-ray-induced unc-22 alleles were positioned within the unc-22 fine-structure map relative to the s8 and s12 alleles (Table 9; Figure 4). The s699 mutation maps between s8 and s18 and is the first allele to be firmly positioned in this region. The s700 mutation maps to the right of s8; however, since this allele failed to recombine with s18, it is not positioned relative to this site.

The sDf19 mutation was isolated as a dominant Unc-22 (twitcher) mutation which was also homozygous lethal (D. Moerman, unpublished results). The fine-structure mapping experiments (Table 9) and complementation results (Table 11) indicate that this mutation is either a deletion which breaks inside the unc-22 gene and extends into let-56, the adjacent gene to the left of unc-22 or an inversion with its breakpoints in the two adjacent genes, unc-22 and let-56. The right endpoint of sDf19 was positioned inside the unc-22 gene between s8 and s18. Since sDf19 affects only unc-22 and let-56, which are probably adjacent genes, it is not possible to distinguish between these two possibilities at this time. The dominant phenotype of sDf19 is presumably due to the breakpoint inside the unc-22 gene.

Table 9

Results of fine-structure mapping experiments between alleles of the unc-22 IV gene in Caenorhabditis elegans

Allele pairs tested*	Total progeny/ heterozygote	Recombinants /total F2 progeny	Map distance**	Left/Right
s12/s32	252	0/351,729	-	-
s18/s35	173	0/155,900	-	-
s18/s36	180	0/274,850	-	-
s8/s699	172	3/341,284	0.0018	s8/s699
s18/s699	222	3/518,092	0.0012	s699/s18
s8/s700	245	8/617,706	0.0026	s8/s700
s18/s700	215	0/786,039	-	-
s8/sDf19	92	0/189,106	-	-
s18/sDf19	133	2/274,600	0.003***	sDf19/s18
s12/sDf19	137	6/274,250	0.009***	sDf19/s12

*The mutation at the left was in the triple: unc-5(e152)unc-22(sx)dpy-4(e1166).

**Map distance was calculated using the formula $d = 2(W)x100/T$ where d = map distance; W = nonUnc-22 recombinants and T = total F2 progeny.

***The number of nonUnc-22 recombinants was multiplied by 4 instead of 2 because of the dominant phenotype of sDf19.

Table 10

Analysis of the exceptional individuals from the fine-structure mapping experiments of the unc-22 IV gene in Caenorhabditis elegans

Alleles tested*	Recombinants	Convertants	Others	Genotype
s12/sDf19	6 Unc-5**			unc-5 + +/unc-5unc-22dpy-4
			1 Unc-5	sterile
s18/sDf19	2 Unc-5			unc-5 + +/unc-5unc-22dpy-4
			1 Dpy-4	+ + dpy-4/unc-5unc-22dpy-4
s8/s699	2 Wild**			+ + dpy-4/+ unc-22 +
	1 Dpy-4			+ + dpy-4/unc-5unc-22dpy-4
			1 Wild	unc-5unc-22 +/ + + +
			2 Unc-5	unc-5 + +/unc-5unc-22dpy-4

Table 10 cont'd

Alleles tested*	Recombinants	Convertants	Others	Genotype
s18/s699	2 Unc-5***			unc-5 + +/unc-22unc-5dpy-4
	1 Wild			unc-5 + +/+ unc-22 +
s8/s700	4 Dpy-4			+ + dpy-4/unc-5unc-22dpy-4
	4 Wild**			+ + dpy-4/+ unc-22 +
		2 Wild		+ + +/+ unc-22 +
		1 Dpy-4 Unc-5		unc-5 + dpy-4/unc-5unc-22dpy-4
			1 Wild	sterile
s18/s700		4 Wild**		+ + +/+ unc-22 +
		2 Dpy-4 Unc-5		unc-5 + dpy-4/unc-5unc-22dpy-4

*The mutation on the left was in the triple: unc-5(e152)unc-22(sx)dpy-4(e1166).

**One of these hermaphrodites was a putative F1 event.

***One of these hermaphrodites was found in the F1 generation.

Figure 4. A fine-structure map of the unc-22 IV gene in Caenorhabditis elegans showing the positions of several alleles and the right breakpoint of sDf19. The positions of the s8, s18 and s12 alleles are from Moerman (1980).

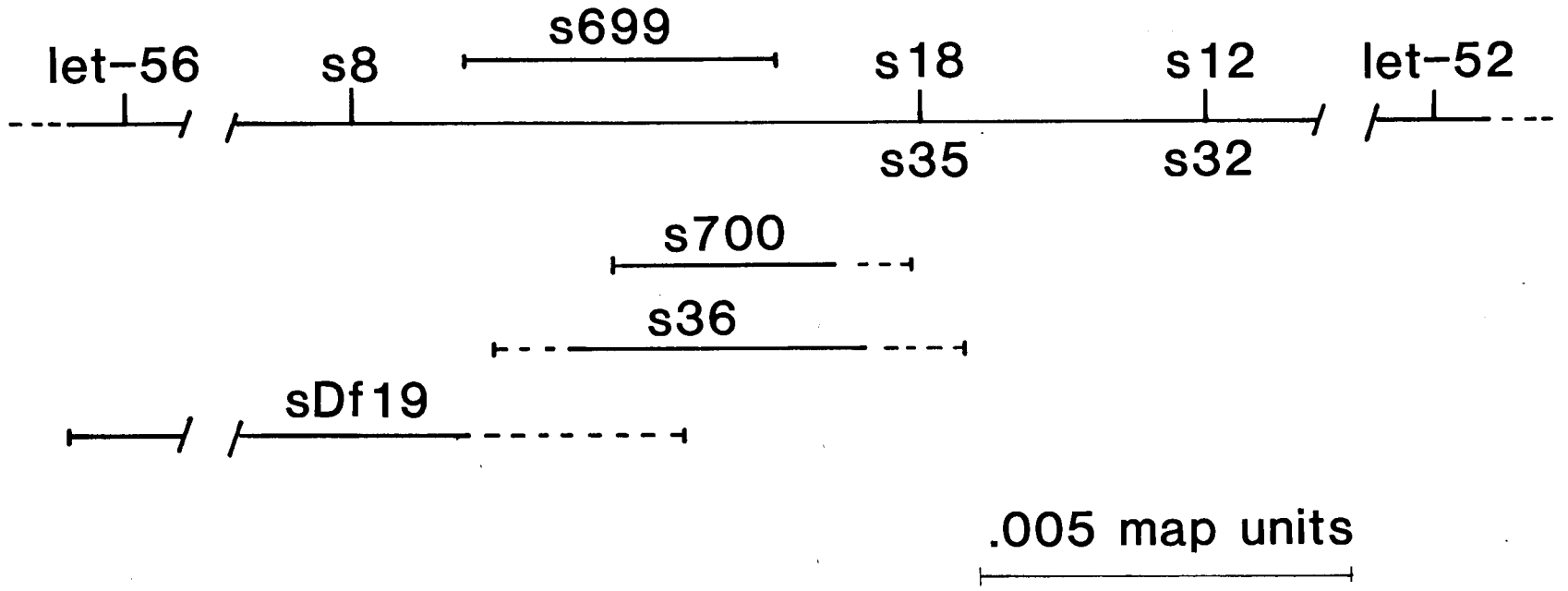

let-56
s8
s699
s18
s12
let-52
s35
s32
s700
s36
sDf19
.005 map units

Table 11

Complementation tests between the sDf19 mutation and six essential genes around unc-22 IV in Caenorhabditis elegans*

Gene(allele)	sDf19
let-52(s42)	+
let-56(s46)	-
let-92(s504)	+
let-60(s59)	+
let-59(s49)	+
let-51(s41)	+

*Complementation is indicated by + and noncomplementation is indicated by -.

VIII. The nT1 translocation as a balancer of unc-22-linked lethal mutations

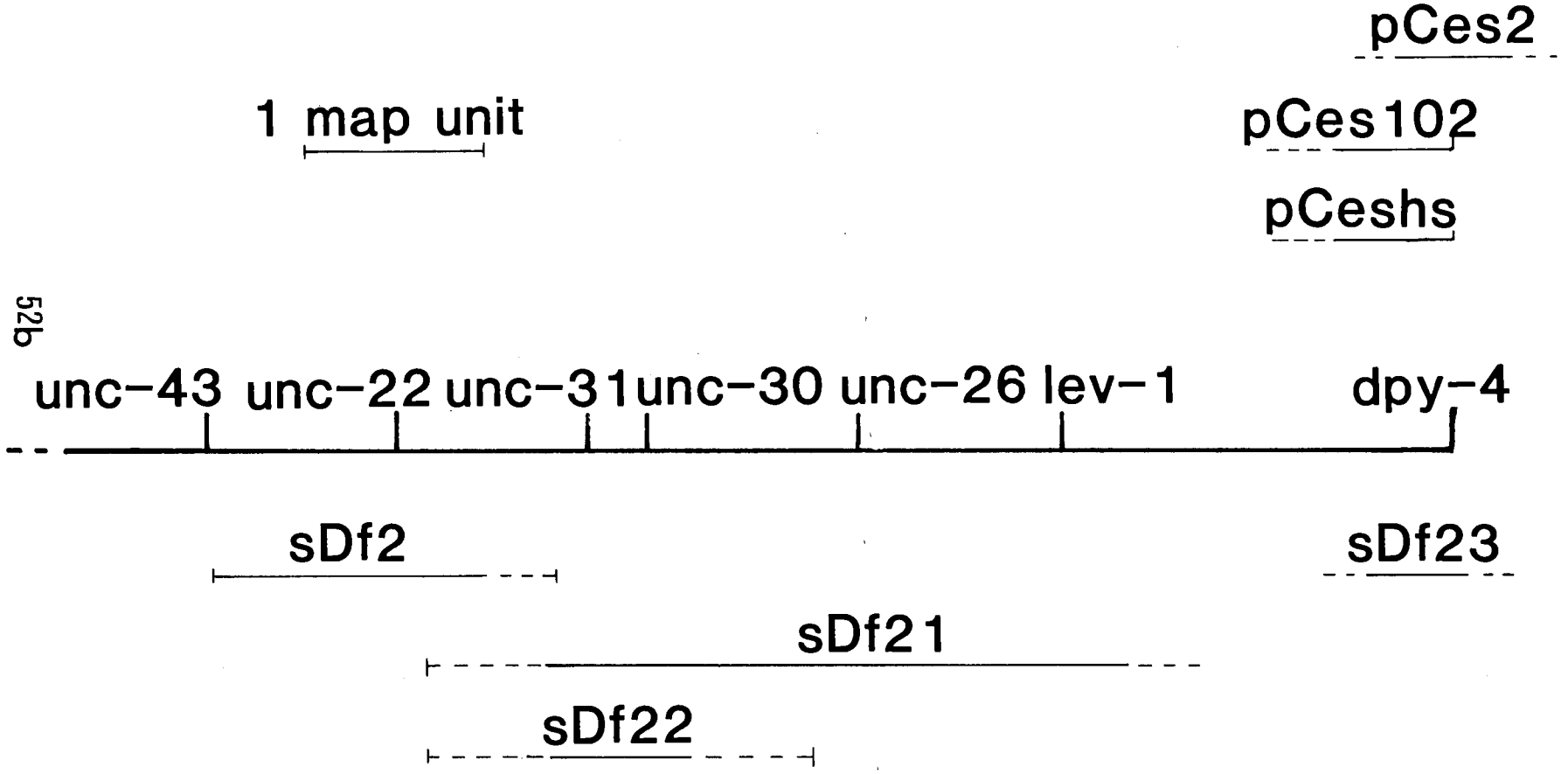
A translocation, nT1(IV;V), between chromosome IV and chromosome V (C. Ferguson, personal communication) was obtained in order to determine if it would be an effective balancer for the unc-22 -linked lethal and sterile mutations isolated in this study. A let-71(s692)unc-22(s7)IV/nT1;+V/nT1 strain was constructed and maintained at 15°C on large plates by selecting single wild-type hermaphrodites at two week intervals. This lethal strain has been maintained for four months without losing the let-71(s692) mutation.

IX. Positioning cloned DNA fragments to Linkage Group IV

Several recombinant plasmids containing C. elegans var. Bristol DNA hybridize to restriction fragments of different lengths in the Bristol (N2) and Bergerac(BO) varieties (Rose et al. 1982). These restriction fragment length differences (RFLD's) between the N2 and BO varieties of C. elegans can be used to position the DNA fragments to the various linkage groups by hybridization to DNA from strains that are homozygous N2 or BO for one particular chromosome and heterozygous N2/BO for the other five chromosomes. Cloned DNA fragments associated with RFLD's were hybridized to DNA from the strains homozygous for

the N2 and BO fourth chromosomes that were constructed in this thesis (see MATERIALS and METHODS) and three of these mapped to this chromosome (Rose et al. 1982; T.Snutch, personal communication). One of the cloned fragments positioned on LG IV contains the structural gene for the 70 kilodalton (Kd) heat shock peptide (Snutch and Baillie 1983; T.Snutch, personal communication). Hybridization to DNA from recombinant strains constructed by myself (MATERIALS and METHODS) and T.Snutch positioned all three cloned fragments near dpy-4 at the right end of LG IV (T.Snutch, personal communication; D.Baillie and K.Beckenbach, personal communication). The positions of the cloned DNA fragments are shown in Figure 5.

Figure 5. A partial genetic map of the right end of linkage group IV in Caenorhabditis elegans. The positions of several deficiencies in the region are shown below the line whereas the positions of several cloned DNA fragments (pCes2, pCes102 and pCeshs) are shown above the line.


pCes2
1 map unit
pCes102
pCeshs
52b
unc-43
unc-22
unc-31
unc-30
unc-26
lev-1
dpy-4
sDf2
sDf23
sDf21
sDf22

D. DISCUSSION

I. Essential genes around unc-22

The region of linkage group IV that was genetically characterized in this thesis is defined by the deficiency, sDf2 (Moerman 1980; Moerman and Baillie 1981). This deficiency includes the unc-22 gene and extends approximately 1.0-map unit to the left of unc-22 and at least 0.5-map units and no more than about 1.1-map units to the right of it. Moerman (1980) had previously identified 14 lethal and sterile mutations that were uncovered by this deficiency and I have identified an additional 20 mutations. All 34 mutations were positioned in this study by recombination mapping and also by complementation mapping with the deficiencies sDf7, sDf8 , sDf9 and sDf10. These mutations identify 20 new essential genes in this region, let-69, let-70, let-71, let-61, let-72, let-59, let-62, let-63, let-64, let-73, let-74, let-91, let-65, let-60, let-92, let-56, let-52, let-66, let-67 and let-68. Four of these genes had previously been identifed by Moerman (1980). Sixteen of the essential genes are in a 1.0-map unit region to the left of unc-22, whereas the remaining four genes are in a region of about 0.5-map units to the right of unc-22. Figure 6 shows that only eight of the essential genes uncovered by sDf2 have more than one allele

(let-59, let-63, let-64, let-72 and let-92 have two alleles each; let-65 and let-68 have three alleles each; let-56 has four alleles). A Poisson analysis of the data indicates that this small region of the C. elegans genome is not saturated for essential genes (see Appendix 3).

Two of the lethal mutations uncovered by sDf2 did not behave as simply as the other lethal mutations. The s681 mutation failed to complement two genes, let-59 and let-62, which map close to each other and may be adjacent genes. This mutation could be a small deletion or some other type of chromosomal rearrangement affecting these two genes. An alternate possibility is that let-59 and let-62 represent a complex locus with complementing alleles (let-59 has two alleles; let-62 has one allele) and s681 fails to complement all other alleles of this gene. Complementing alleles have been observed for the unc-15 I locus in C. elegans (Rose 1980) and for the maroon-like locus in D. melanogaster (Finnerty 1976).

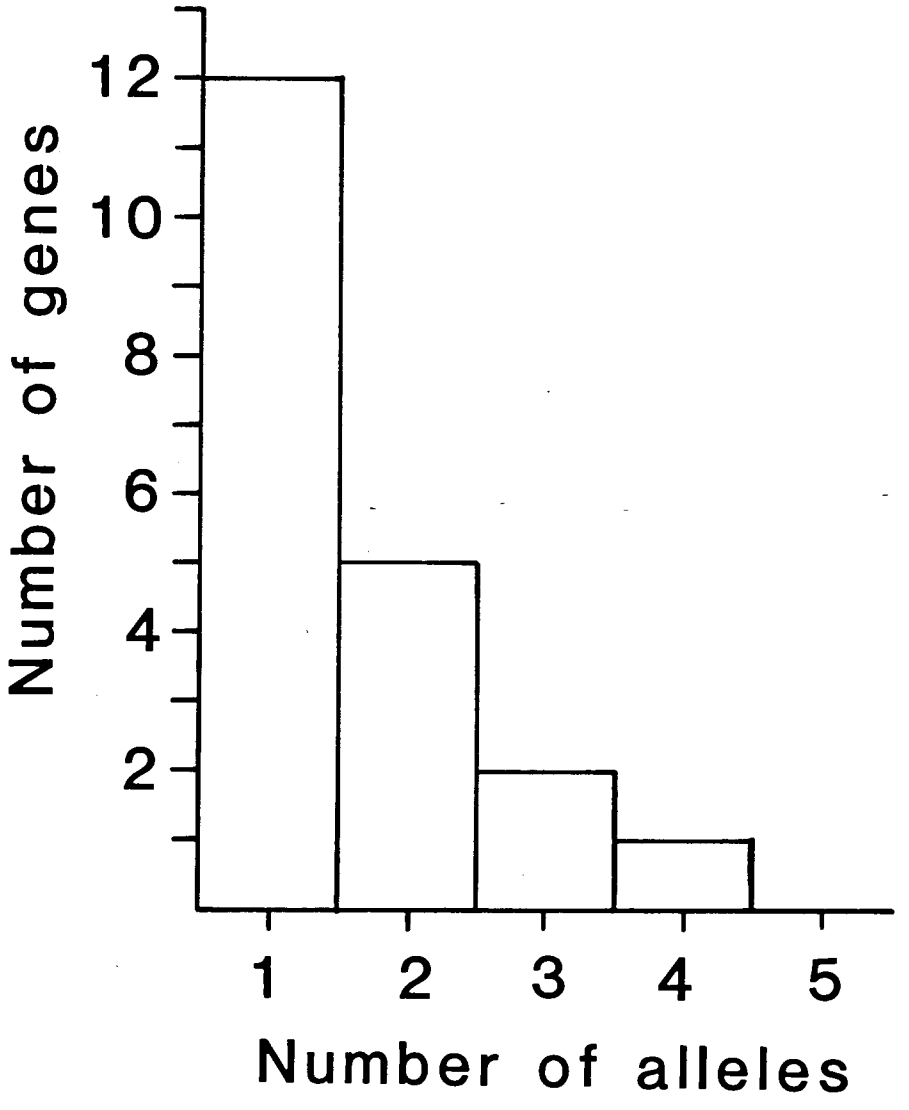
In the case of s698, the recombination mapping and deficiency mapping experiments gave conflicting results. Recombination mapping placed s698 0.04-map units to the right of unc-22 (Table 3). These results suggest that s698 is very close to unc-22. In contrast, s698 complemented the four deficiencies, sDf7, sDf8, sDf9 and sDf10, as well as let-66(s176) and let-67(s214) but failed to complement any of the three alleles of let-68. Therefore, based on deficiency mapping and complementation analysis, this mutation lies in zone 6 and is an

allele of let-68 which maps 0.5-map units to the right of unc-22.

An explanation for these conflicting data is that s698 is an inversion with one breakpoint in the let-68 gene and the other breakpoint near unc-22 in a nonessential region, or in another essential gene in zone 6 which has not yet been identified. If s698 is an inversion, then it should suppress crossing over in the region between unc-22 and unc-31. Two additional mutations which lie outside the sDf2 deficiency gave results similar to s698. Recombinationally, s212 and s690 map close to unc-22, 0.2-map units to the left and 0.08-map units to the right, respectively, yet they complement sDf2.

All of the lethal and sterile mutations that were positioned in this study were isolated after mutagenesis treatment with ethyl methanesulfonate (EMS). This mutagen has been shown to cause, primarily, G:C to A:T transitions (Coulondre and Miller 1977; Karn et al. 1983) and is thought to act by adding an ethyl group to the N7 position of guanine causing it to behave as a base analogue of adenine and pair with thymine (Strickberger 1968). Therefore, the mutations used in this analysis should be point mutations. The majority of these mutations do appear to be point mutations.

Figure 6. A bar graph indicating the number of essential genes in the sDf2 IV region of Caenorhabditis elegans represented by one or more mutations.


Number of genes
12
10
8
6
4
2
1
2
3
4
5
Number of alleles

II. Nonessential genes around unc-22

Four genes with visible phenotypes, unc-43, unc-31, dpy-20 and unc-30, which had been previously positioned near unc-22 were positioned relative to this gene (see Figure 1; Table 5) and tested for complementation with sDf2. The dpy-20(e1282) mutation failed to complement sDf2 (Table 6) and, therefore, lies inside the region uncovered by this deficiency while the three remaining genes lie outside the sDf2 region. The dpy-20 gene was further positioned into zone 2 by complementation mapping with the sDf7, sDf8, sDf9 and sDf10 deficiencies (Table 6). Another nonessential gene, dpy-26, which has not been firmly positioned, also failed to complement sDf2. The homozygous dpy-26 mutant is inviable (J.Hodgkin, personal communication), thus this mutation could be an allele of one of the essential genes in the region. At present there are, at least, two nonessential genes, dpy-20 and unc-22, in the sDf2 region in addition to the 20 essential genes.

III. The size of the deficiency, sDf2

The extent of sDf2 was determined from the recombination data obtained for the lethal and sterile mutations. If one considers only the mutations uncovered by sDf2, then the left endpoint of this deficiency is 1.2-map units from unc-22.

However, if the mutations outside sDf2 are considered then sDf2 may break as close as 0.8-map units from unc-22. These results indicate that the left endpoint of sDf2 is between 0.8- to 1.2-map units or approximately 1.0-map unit from unc-22. The right end point of sDf2 is less firmly positioned. sDf2 includes let-68 which maps 0.5-map units from unc-22 but does not include unc-31, which is approximately 1.1-map units from unc-22. Therefore, sDf2 extends at least 0.5-map units and possibly as far as 1.1-map units to the right of unc-22. The fact that several of the lethal mutations positioned outside sDf2 map closer to unc-22 than some of the lethals positioned inside this deficiency shows that map distances alone cannot be used to position close genes relative to one another. These results also illustrate the importance of deficiencies for the left-right positioning of very close genes.

IV. Organization of essential genes in the unc-22 IV region

The asymmetric distribution of the 20 lethal genes around unc-22 is not surprising if the right endpoint of sDf2 is only 0.5-map units from unc-22. However, as previously indicated, the right endpoint of sDf2 is not firmly positioned and could extend as far as 1.1-map units from unc-22. In that case the data may reflect a clustering of lethal sites. Brenner (1974) noted that the visible genes tend to map in a cluster on the autosomal linkage groups and those of LG IV have been found to cluster on

its right arm with the center of the cluster to the left of unc-22 (Herman and Horvitz 1980). Herman (1978) found that on LG II lethals cluster in the same region as the visibles.

All but one of the nineteen lethal mutations which lie outside the sDf2 region, including s41, s43, s44 and s45 (see Moerman 1980), map to the left of unc-22 (Table 4). The single mutation, s690, which maps to the right of unc-22 behaves anomalously and cannot be accurately positioned (see above). The present results indicate that the lethal genes around unc-22 are clustered to the left of this gene. There are at least 16 essential genes in the 1.0-map unit interval to the left of unc-22 and only four essential genes in the 1.0-map unit interval to its right.

An interesting aspect of this analysis is that no lethal mutations, either inside or outside sDf2, have been found in the interval between let-68 and unc-31 which are 0.5- and 1.1-map units, respectively, to the right of unc-22. As yet, no nonessential genes have been positioned in this region. The two genes, dpy-26 and lin-24, lie between unc-22 and unc-31 (J.Hodgkin, personal communication; C.Ferguson, personal communication) but they have not been positioned with respect to let-68. dpy-26 is uncovered by sDf2, but lin-24 has not been tested with this deficiency. This region may contain relatively few genes, or if there are genes in this region they may be cryptic genes or members of multigene families.

V. Genomic organization in Caenorhabditis elegans

In addition to the unc-22 region, two other genomic regions of C. elegans have been characterized. Meneely and Herman (1979; 1981) found 24 essential and eight nonessential genes in a 7.5-map unit interval at the tip of the X chromosome, whereas Rose and Baillie (1980) found seven essential and five nonessential genes in a 0.7-map unit interval around unc-15 I. Since none of these regions are completely characterized yet, it is necessary to estimate the total number of genes in each region before they can be compared. This has been done for the X chromosome region by Meneely and Herman (1979; 1981). These authors calculated that there are a total of 45 essential genes in this region based on the fraction of mutated genes in their sample that are represented by more than one mutant. Using this same method I estimate that the total number of genes in the unc-15 I and unc-22 IV regions are 15 and 32, respectively (see Appendix 3). These data indicate a difference in gene density between the small, autosomal regions (28 and 22 genes/map unit, respectively) and the larger region of the X chromosome (7 genes/map unit). It will be interesting to see whether this difference is due to the regions being studied, to the chromosomes being studied or to a combination of both.

VI. The DNA content of the sDf2 region

The amount of DNA in the sDf2 region of linkage group IV can be estimated from the haploid genome size of C. elegans and the size of the genetic map. The present C. elegans genetic map consists of 300 map units (M.M. Swanson, personal communication) whereas the DNA content of this organism is 8×10^7 base pairs (Sulston and Brenner 1974). Assuming that the recombination rate per base pair of DNA is equal in all regions of the genome, there are approximately 267,000 base pairs of DNA/map unit in C. elegans. The twenty essential genes defined in this thesis lie in an interval of 1.5-map units or 400,000 base pairs of DNA. Using the estimate of 32, as the number of essential genes uncovered by sDf2 (see above) and including the nonessential genes, unc-22 and dpy-20, there are roughly 34 genes in the 400,000 bp region uncovered by sDf2. This translates into about 11,800 bp of DNA/gene which is comparable to the estimated value of 20,000 bp/gene in Drosophila melanogaster (Alberts et al. 1982). Using the estimate of 11,800 bp of DNA/gene as an average value for all C. elegans genes then there is sufficient single copy DNA in C. elegans to code for 5700 genes. This estimate of the total number of genes in the nematode is much larger than Brenners original estimate of 2000 genes (Brenner 1974) which was based on mutation frequency. It is important to note that the estimate of 5700 genes assumes that the sDf2 region is an average region in terms of gene density, a fact that is probably

not true.

VII. Characterization of the lethal and sterile mutations

The effective lethal phases of twenty-six of the lethal mutations (19 inside sDf2 and 7 outside sDf2) that were positioned in this thesis had previously been determined by Moerman (1980). The remaining 27 mutations were characterized in this study (Tables 7 and 8). Both the range and distribution of the lethal phases displayed by these two sets of mutations are very similar. The mutations uncovered by sDf2 block at various developmental stages, from first stage larva to sterile adult. However, none of these 34 mutations arrest development during embryogenesis. From these results it would appear that the genes in the sDf2 region are required at all stages of development except embryogenesis. However, the absence of embryonic lethals may be due to the strong maternal effects in C. elegans. I did not expect to find any embryonic lethals in the region uncovered by sDf7, sDf8, sDf9 or sDf10, since strains homozygous for these deficiencies survived embryogenesis. However, there may be genes essential for embryonic development that have not yet been identified in the regions uncovered solely by sDf2, as this deficiency blocks during embryogenesis when homozygous. Alternatively, the fact that sDf2 homozygotes are embryonic lethals may be a result of the large size of this deficiency.

As indicated earlier, eight of the 20 essential genes identified in this analysis have more than one allele each. Both alleles of let-59, let-64 and let-92 have the same effective lethal phases. However, this is not the case for the other five genes (Table 7; Moerman 1980). The let-63 and let-72 genes have two alleles, one blocking at the mid-larval stage and the other allele blocking at the late larval stage of development. Two of the three alleles of let-65 arrest development at the mid-larval stage whereas the third allele arrests development as a sterile adult. One of the let-68 alleles is a late larval lethal while the two other alleles arrest development as sterile adults. Finally, two of the four let-56 alleles are mid-larval lethals while the two remaining alleles are late larval lethals. Except for let-65, the terminal phenotypes of the various alleles of these genes do not differ by more than one developmental stage. The alleles of these five genes which arrest development at the later larval stages are probably less severe mutations than those blocking at the earlier larval stages. For example, a missense mutation resulting in a gene product with partial activity would be a less severe mutation than a null mutation where no gene product is made. Perhaps, some of the later alleles produce a gene product with enough residual function to allow hermaphrodites that are homozygous for these mutations to develop to a later larval stage than hermaphrodites that are homozygous for the earlier alleles which do not have any gene activity. It is not known whether any of the mutations

positioned in this study are null alleles. This could be determined by testing the mutations for suppression by sup-7 X, an informational suppressor isolated by Waterston (1981).

VIII. The genes adjacent to unc-22

The initial analysis of eight lethal mutations in the unc-22 region (Moerman 1980) identified two genes, let-56 and let-52, as the flanking genes to the left and right of unc-22. The data obtained in this analysis, which characterized a further 32 lethal mutations, still support this conclusion. let-56 maps 0.03-map units to the left of unc-22 in zone 4 while let-52 maps 0.07-map units to the right in zone 5 (Tables 3 and 6; Figure 2). Three additional alleles of let-56, s52, s168 and s173, were identified in this study. Since these two genes map so close to unc-22 they could possibly be lethal alleles of this gene. I have demonstrated that let-52(s42) is not an allele of unc-22 because it can be separated from the unc-22 locus by the deficiency, sDf8 (Table 6).

IX. The size of the unc-22 gene

The fine-structure map of the unc-22 gene constructed by Moerman (1980) consisted of 15 unc-22 alleles and the estimated size of this gene was between 0.01- and 0.02-map units. I have positioned two new alleles into this map. The s699 mutation was

positioned between s8 and s18 and is the first mutation to be firmly positioned in this region. The s700 allele mapped to the right of s8 but was inseparable from s18. Since both of these mutations map within the existing boundaries of the unc-22 fine-structure map the size of the gene has not been increased.

In addition to positioning these new alleles, I have also refined the positions of 3 of the original 15 unc-22 alleles. The s32 allele was positioned at the right boundary of the unc-22 fine-structure map on top of the s12 and e66 alleles while the s35 and s36 mutations were found to be inseparable from the s18 site. Since s32 is suppressed by the informational suppressors, sup-5 III and sup-7 X (Moerman 1980; D.Moerman, personal communication) it is, therefore, a mutation in the coding element of the unc-22 gene. The only other mutation that is known to be in the coding element of this gene is the dominant allele, m52, which has been positioned close to s18 (Moerman 1980). Therefore, the coding element of the unc-22 gene includes, at least, the region of the fine-structure map from the right boundary to the m52 allele (Figure 4).

The adjacent genes on either side of unc-22, let-56 and let-52, map 0.03- and 0.07-map units, respectively, from the s7 allele of unc-22 (Table 3). The distance between these two genes, about 0.1-map units or 26 kilobases (Kb) of DNA (see above), represents the maximum size of the unc-22 gene. The existing fine-structure map of unc-22 occupies only a small portion of this interval. The coding element of the unc-54 gene

in C. elegans is 7,266 bp in length (Karn et al. 1983) while the three genes in the ovalbumin cluster in chicken are 6-8 Kb in length with intergene regions of 5-10 Kb (Royal et al. 1979). If the unc-22 gene is similar to the rosy locus in Drosophila melanogaster (Chovnick et al. 1977; McCarron et al. 1979) then most, if not all, of the alleles positioned in the present fine-structure map will be mutations in the coding element of this gene. There is certainly sufficient DNA for the regulatory element of the unc-22 gene to be located outside the existing boundaries of the fine-structure map.

X. Internal deletions of the unc-22 gene

The recombination data obtained from the mapping experiments using s699 and s700 (Table 9) suggest that the distance between the s8 and s18 alleles of the unc-22 gene is approximately .003-map units. This value is considerably less than the .007-map units obtained by Moerman (1980) for the same interval. The shorter map distance obtained in this thesis could be explained if the s699 and s700 mutations are small deletions or inversions. In that case, s699 would delete an interval of about 0.004-map units between the s8 and s18 sites. The exact size and position of s700 cannot be determined from the existing data. It would not be surprising if these two mutations were small deletions or inversions since both of them were induced with gamma-radiation (R.Rosenbluth, personal comunication). The

other, obvious explanation is that the map distances determined in this analysis are not comparable to those determined by Moerman (1980). This could be due to the different methods employed to calculate the number of F2 progeny or, more likely, to a difference in ability to identify the nonUnc-22 individuals against the background of twitching hermaphrodites.

The fine-structure mapping data of Moerman (1980) tentatively positioned the formaldehyde-induced s35 and s36 alleles of unc-22 in the region between s8 and s18. In this thesis these two mutations were mapped against the s18 allele to confirm these positions. Since both mutations failed to separate from the s18 mutation they could not be positioned relative to it. However, the results do indicate that the sites of the s35 and s36 mutations must be close to the s18 site. This is consistent with the previously determined position of s35 but is not in agreement with that of s36. Although Moerman (1980) placed the s36 mutation about 0.0026-map units to the right of s8 on the fine-structure map, his data suggested that this mutation was an internal deletion of unc-22. My results support this suggestion, indicating that s36 extends close to and may even delete the s18 site. Formaldehyde has been shown to induce deletions in C. elegans (Moerman 1980; Moerman and Baillie 1981) but it is not known whether this mutagen also induces point mutations. Although the s35 mutation behaves as if it were a point mutation it could be a small deletion.

XI. Analysis of a dominant Unc-22 mutation

The other mutation affecting the unc-22 gene that was positioned in this analysis was sDf19, which has a dominant Unc-22 (Twitcher) phenotype and is also homozygous lethal. Both the fine-structure recombination data (Table 9) and the results of the complementation analysis (Table 11) indicate that this mutation is a deficiency or inversion affecting the two adjacent genes let-56 and unc-22. The dominant twitcher phenotype of sDf19 is, presumably, a result of the deficiency or inversion endpoint which has been positioned inside the unc-22 gene (Figure 4) while the recessive lethal phenotype is either the result of another breakpoint in let-56 or a deletion of this gene.

The dominant Unc-22 phenotype of sDf19 indicates that an aberrant unc-22 protein is being synthesized from the mutant gene. If, in fact, this mutation is a deletion of the left end of the unc-22 gene, then in order for this gene to be transcribed, the unc-22 promoter must be located at the right end of the gene. This implies that the direction of transcription of the unc-22 gene is from right to left on the genetic map. Another, although more complicated, explanation of the synthesis of a mutant unc-22 protein is that the unc-22 promoter has been deleted and the remaining segment of the gene is now coupled to another promoter in the same reading frame and is transcribed from this new promoter.

XII. Analysis of exceptional individuals

In four of the fine-structure mapping experiments that were done, no nonUnc-22 progeny were observed whereas a total of 37 exceptional individuals were found in the remaining six mapping experiments. All of the 37 nonUnc-22 hermaphrodites were progeny tested to determine their genotype (Table 10). Twenty-two of these individuals were the result of a recombination event between the two unc-22 alleles being mapped, and nine were the result of a conversion event of one of the unc-22 alleles. Two of the remaining six individuals were sterile and the other four probably resulted from two independent events. For example, the two anomalous unc-5 + +/unc-5unc-22 + individuals observed in the experiment to position s699 relative to s8 could be the result of a crossover between unc-22(s8) and dpy-4 to generate an unc-5unc-22(s8) + F1 chromosome. A conversion of the s8 allele on this chromosome to a wild-type allele in an F1 individual with the genotype unc-5unc-22(s8) +/+ s699 + would result in an Unc-5 F2 individual if the unc-5 + + chromosome paired with an unc-5 unc-22 + chromosome. However, the exact same result would be obtained if the unc-5 marker crossed onto the + s699 + chromosome to produce an unc-5unc-22(s699) +/unc-5unc-22(s8)dpy-4 F1 individual. Conversion of the s699 mutation and the subsequent pairing of the resulting unc-5 + + chromosome with and unc-5unc-22(s699) + chromosome would result

in the Unc-5 hermaphrodite that was observed.

The only confirmed conversion events of unc-22 mutations were found in the mapping experiments with s700. Eight recombinants and three convertants were observed in the experiment with s8 whereas all six exceptional progeny in the experiment with s18 were convertants. The rates of these conversion events were 3 out of 617,706 (1/205,000) progeny in the s8/s700 experiment and 6 out of 786,039 (1/131,000) progeny in the s18/s700 experiment. The conversion event is not specific to the s700 mutation since both s8 and s18 also converted, although in both cases, the s700 allele converted twice as often as the other two alleles. These results do not exclude the possibility that the s700 mutation is a deficiency since Fink and Styles (1974) working on the HIS-4 locus in yeast observed conversion of deletion mutations. However, if s700 is a deletion it cannot include the s18 mutant site.

Gene conversion has been observed in several organisms including C. elegans (Rose 1980; Moerman 1980). Moerman (1980) observed 104 nonUnc-22 individuals in his unc-22 fine-structure mapping experiments, 15 of which were the result of a conversion of one of the unc-22 mutations to wild-type. Clearly, the ratio of recombinants to convertants obtained by Moerman (87:15) is much larger than the 22:9 ratio obtained in these experiments. This result can be explained by the fact that the unc-22 alleles that were positioned in this study were generally much closer than those mapped by Moerman (1980). Chovnick et al. (1971)

noted that in D. melanogaster the number of crossover events compared to conversion events increased with increased distance between the two alleles being mapped.

XIII. Unusual recombination events

An unusual result of the fine-structure mapping experiments was the occurrence of one confirmed F1 recombination event and four putative F1 events (Table 10). One exceptional individual was identified in the F1 generation and the genotype of this hermaphrodite was consistent with a recombination event having occurred between the s699 and s18 alleles of the unc-22 gene. The four putative F1 events were identified in the F2 generation by the presence of a large number (>20) of nonUnc-22 progeny on a single plate. Three of these events were recombination events and the other one was a conversion of s700.

Moerman (1980) also observed bursts of nonUnc-22 progeny in his unc-22 fine-structure mapping experiments. Since the burst sizes were smaller than expected if these individuals were the progeny of an F1 recombinant or convertant Moerman suggested that these exceptional individuals might result from a mitotic recombination or conversion event during gametogenesis. F1 recombination events do occur since one was observed in this analysis. The frequency of this event was about 1/93,000 progeny whereas the frequencies of the F2 events in the same experiment were about 1/100,000 progeny. Whether the other, putative F1

recombinants and convertant observed in this study actually resulted from an F1 event or from some other event is not known. Bursts of exceptional progeny in fine-structure mapping experiments are not limited to the unc-22 gene; this phenomenon has also been observed for the unc-15 gene in C. elegans (Rose 1980) and the rosy gene in D. melanogaster (Chovnick et al. 1971).

XIV. Resolution of the fine-structure mapping procedure

The closest unc-22 alleles that were separated in this study were s699 and s18, which were found to be approximately 0.0012-map units apart. Using the estimate of 267 Kb of DNA per map unit that was calculated earlier, these two mutant sites are separated by about 320 nucleotide pairs. Moerman (1980) separated two alleles, s16 and s8, that were 0.0007-map units or about 200 base pairs apart. Just over half a million F2 progeny were screened in the experiment that positioned s699 relative to s18 (Table 9). The largest sample that was screened in this analysis consisted of about 786,000 F2 hermaphrodites (Table 9). With a sample size this large it should be possible to separate mutation sites that are only 70 nucleotides apart.

In addition to the unc-22 gene, fine-structure maps have been constructed for three other genes in C. elegans: unc-13 and unc-15 (Rose and Baillie 1980) and unc-54 (Waterston et al. 1982). All three of these genes are on linkage group I (see the

genetic map in Appendix 1). unc-13 and unc-15 are adjacent genes and they lie in a region that has been characterized genetically (Rose and Baillie 1980). However, there are only a limited number of mutations in these two genes. The unc-15 gene is the structural gene for the muscle protein, paramyosin, whereas the unc-13 gene product has not been identified. The unc-54 gene codes for the major myosin heavy chain found in the body-wall musculature of C. elegans (McLeod et al. 1977). The fine-structure map of this locus is very extensive and, in addition, unc-54 has been cloned (McLeod et al. 1981) and sequenced (Karn et al. 1983). The region around unc-54 is not well characterized nor have the adjacent genes to the left and right of this gene been identified. The unc-22 gene is located on a different chromosome from the other three genes. Although the fine-structure map of the unc-22 gene is not as extensive as the unc-54 map, it can easily be expanded since there are many alleles of this gene. The region around unc-22 is well characterized and the two adjacent genes have been identified.

XV. The unc-22 region as a model system

This thesis establishes the unc-22 IV gene and the adjacent region in C. elegans as an experimental system for studying the regulation of a muscle gene and the organization of the genome around this gene. The unc-22 region is, at present, the best genetically characterized region of the C. elegans genome. C.

elegans is the only other higher eucaryote, besides D. melanogaster, where this type of genetic analysis has been undertaken.

As an experimental system for studying genomic organization and gene regulation, the unc-22 region compares very favorably to the rosy system in D. melanogaster. A fine-structure map of the unc-22 gene exists (Moerman and Baillie 1979; Moerman 1980; this thesis) and it is possible to identify mutations in the coding element of this gene using informational suppressors (Waterston 1981). The unc-22 gene product, although not yet identified, is a muscle protein and, presumably, is found only in muscle cells. The adjacent genes on both sides of the unc-22 gene have been identified, thus delimiting the extent of this gene. Many of the essential genes in the region around unc-22 have been identified, making it possible to estimate the gene density in this region and, in combination with studies of other regions, to estimate the total number of genes in C. elegans. The approximate time that these genes are required during development has also been determined. It should be possible to study the genes in the unc-22 region to determine whether they are functionally related. Finally, the unc-22 gene and the surrounding region are both amenable to a molecular analysis.

XVI. Limitations of experimental techniques

In this thesis, I have concentrated on identifying and positioning essential genes in the sDf2 region, assuming that the majority of the genes in C. elegans are essential for development. The method employed (i.e. isolating point mutations in essential loci) will not identify genes that are present in many copies in the genome or genes that have cryptic or unusual phenotypes. The nonessential genes that were positioned in this region had been previously isolated. No attempt was made in this analysis to specifically isolate mutations with visible phenotypes. It is, therefore, possible that there are more than two nonessential genes in the sDf2 region.

Two methods were used to position the mutations representing the essential and nonessential genes in the sDf2 region; recombination mapping and deficiency mapping. Deficiency mapping is a more accurate method of positioning mutations assuming that there is nothing unusual about the deficiency. The accuracy of recombination mapping depends on several factors, including the number of F1 progeny sampled, the number of recombinant individuals obtained and ability to identify recombinant individuals. Large numbers of F1 progeny were counted to determine the map distances obtained in this study, therefore, these distances should be accurate. The recombinants used to calculate these map distances were real since all exceptional individuals were confirmed by progeny testing. The

map distances obtained would underestimate the actual distances if some recombinants were not identified. However, I am confident that most, if not all, of the recombinants were identified. The fact that the positions obtained by recombination mapping and deficiency mapping were similar for most of the mutations indicates that the genes in this region are accurately positioned.

The number of essential genes defined by the the lethal mutations uncovered by sDf2 was determined by complementation analysis. This analysis assumes that each mutation represents a single gene and, also, that none of the mutations are complementing alleles of the same gene. The presence of complementing alleles in this sample would reduce the number of genes identified.

The calculation used to estimate the total number of essential genes in the sDf2 region (Appendix 3) assumes that all of the genes in this region are equally mutable. This is probably not true since mutability would be affected by gene size, GC content and the presence of regions in the gene where mutations do not affect gene function.

The distances between alleles of the unc-22 gene that were obtained may not be very accurate since the total sample size was only estimated and, in most cases, there were only a few recombinants observed. In addition, some of the nonUnc-22 recombinants may have been missed.

XVII. Summary and Prospects

As indicated in the INTRODUCTION the aim of this analysis was to establish a genetically well defined region of the C. elegans genome in order to study the organization of the genetic material in higher eucaryotes. The sDf2 region described in this thesis is at present the most extensively characterized region in this organism. Twenty essential genes and two nonessential genes have been identifed and positioned in an interval of 1.5-map units on LG IV. These genes were positioned by two independent methods: recombination mapping and deficiency mapping. With very few exceptions, the positions obtained by these two mapping methods were in very good agreement. The genes in this region appear to be required at all stages of development except embryogenesis. The present data indicate that the twenty essential genes identified in this study represent about 2/3 of the essential genes in this region. Therefore, assuming a total of 32 essential and two nonessential genes, the genes in this interval are separated by approximately 0.05-map units.

This genetic analysis of the sDf2 region complements the fine-structure analysis of unc-22 IV, one of the genes in this region (Moerman and Baillie 1979; Moerman 1980; this thesis). The maximum size of this gene, which is defined as the distance between the adjacent genes on either side of unc-22, is 0.1-map units while the minimum size of this gene, which is represented

by the present fine-structure map, is 0.02-map units. At present, the coding element of the unc-22 gene includes the right end of the fine-structure map. Three possible internal deletions of unc-22 have been identified and positioned. In addition, the right endpoint of a deficiency or an inversion affecting the adjacent genes, let-56 and unc-22, has been positioned inside the unc-22 gene. This formaldehyde-induced mutation is associated with a dominant Unc-22 phenotype indicating that the mutant unc-22 gene is transcribed. The simplest interpretation of these data is that the left end of the unc-22 gene has been deleted and that the unc-22 gene is, therefore, transcribed from right to left on the genetic map. This hypothesis will have to be confirmed by a molecular analysis.

The 70 kilodalton (Kd) heat shock gene in C. elegans (Snutch and Baillie 1983) has been cloned and positioned near dpy-4 on linkage group IV using a restriction fragment length difference associated with the cloned DNA between the Bristol and Bergerac strains of C. elegans (see Results; T. Snutch, personal communication). Two additional recombinant DNA clones have been mapped in the same region (Rose et al. 1982). Attempts to clone the unc-22 gene by chromosome walking from these sites, which are approximately 5-map units from unc-22 are in progress (D.Baillie and K. Beckenbach, personal communication). Several deletions have been isolated in the region between unc-31 and dpy-4 on linkage group IV (Figure 5) as part of a genetic

characterization of that region (D.Baillie and G.Wild, personal communication). These deficiencies will facilitate the chromosome walking by reducing the length of the region between the cloned DNA fragments and the unc-22 gene. The sDf2, sDf7, sDf8, sDf9 and sDf10 deficiencies will be extremely useful in identifying recombinant DNA clones which contain DNA in the region adjacent to unc-22 while the internal deletions of the unc-22 gene will identify clones containing the unc-22 gene.

The molecular analysis of the unc-22 gene will enable us to determine the nature of the unc-22 mutations, particularly sDf19, and to compare the positions of these mutations in the fine-structure map with their positions in the DNA sequence. Since the cloning of the unc-22 gene should make the identification of its protein product relatively easy, the effects of these mutations on this protein can also be studied. In addition, a molecular analysis will allow for the direction of transcription of this gene to be elucidated. DNA sequencing and mRNA hybridization experiments will define the unc-22 structural element. It will then be possible to determine if any of the mutations in the fine-structure map lie outside this region. Once the structural element has been defined it would be very useful to isolate and position regulatory mutants. A similar molecular analysis of the entire region uncovered by sDf2, along with the genetic analysis, should determine the exact number of genes in this region, the amount of DNA required for each gene and the relationship between map distance and

amount of DNA in this region of the C. elegans genome.

XVIII. Proposals for further research

Listed below are several experiments that could be done to extend the genetic analysis presented in this thesis.

1. Generate additional deficiencies in the unc-22 region using formaldehyde and gamma-radiation in order to left-right position the essential genes in this region.

2. Continue isolating lethal and sterile mutations in the sDf2 region until all of the essential genes in this region have been identified.

3. Identify null mutations of the essential genes by testing the lethal and sterile mutations which define these genes with the informational suppressor sup-7 X (Waterston 1981).

4. Determine if the sDf8 deficiency breaks inside the unc-22 gene by fine-structure mapping with the s12 allele of unc-22.

5. Determine if any of the 50 unmapped unc-22 alleles are suppressable by sup-7 and position any that are suppressed into the fine-structure map to define the coding element of this gene.

6. Identify regulatory mutants of the unc-22 gene and position these mutations into the fine-structure map.

7. Determine whether rec-1, a mutation which increases recombination frequency in C. elegans (Rose and Baillie 1979), increases the frequency of intra-geneic recombination in the unc-22 gene.

8. Analyze the sDf2 region at the molecular level by cloning and sequencing all of the DNA in this region.

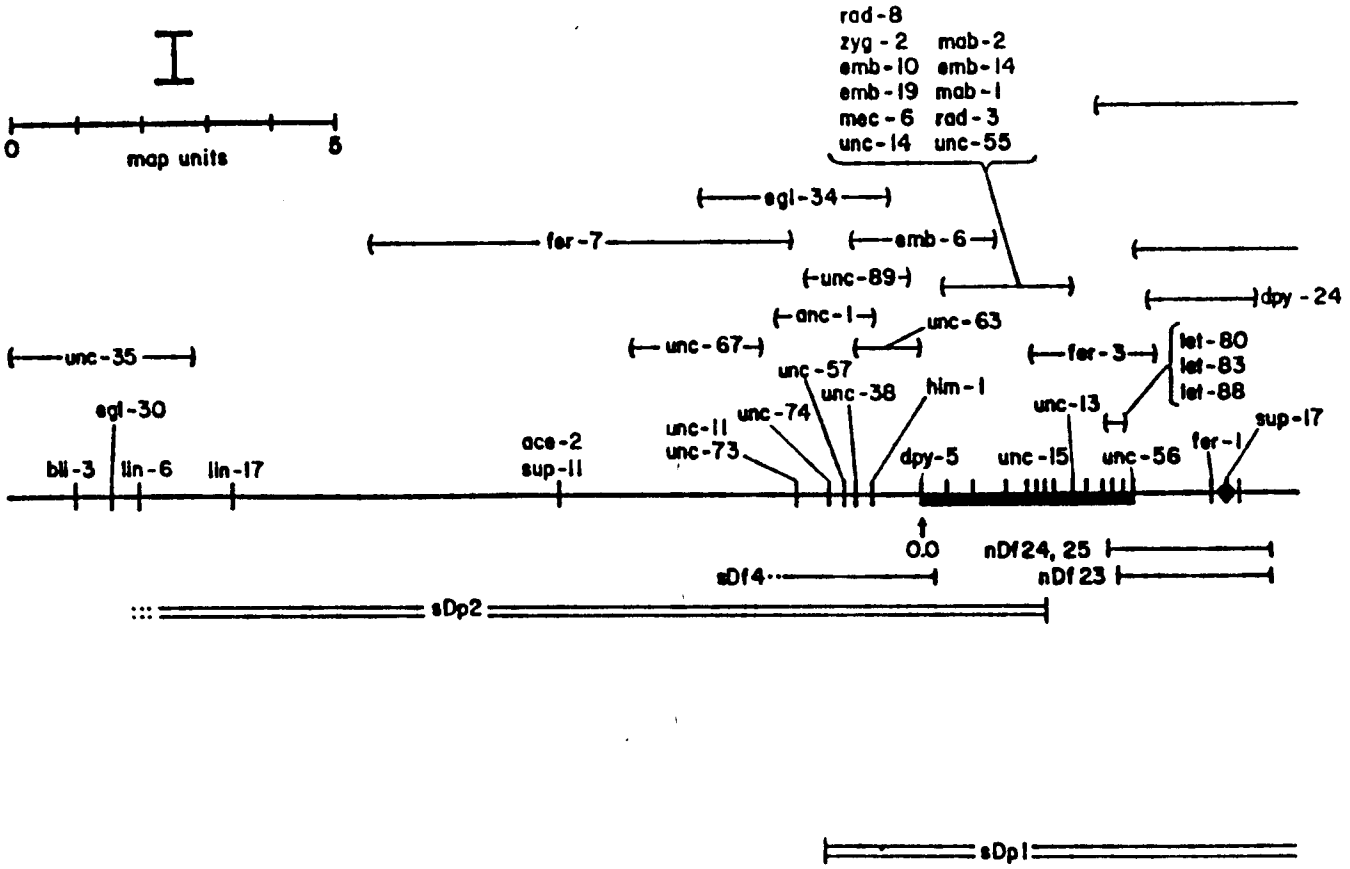
9. Detemine the protein products of unc-22 and the surrounding genes by hybrid selection using recombinant DNA clones.

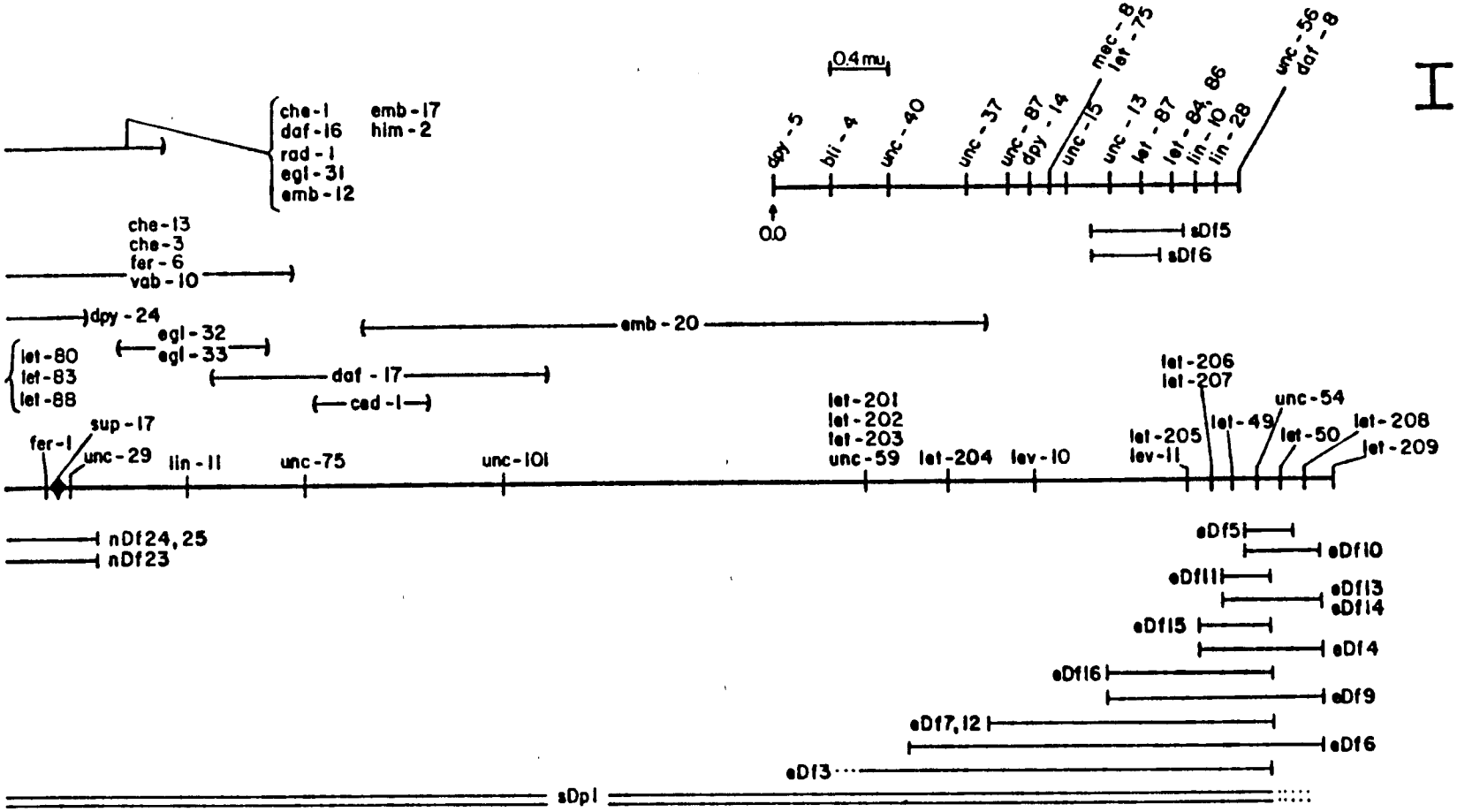
10. Clone and sequence some of the mutant unc-22 genes, particularly the formaldehyde-induced mutants, in order to determine the nature of these lesions.

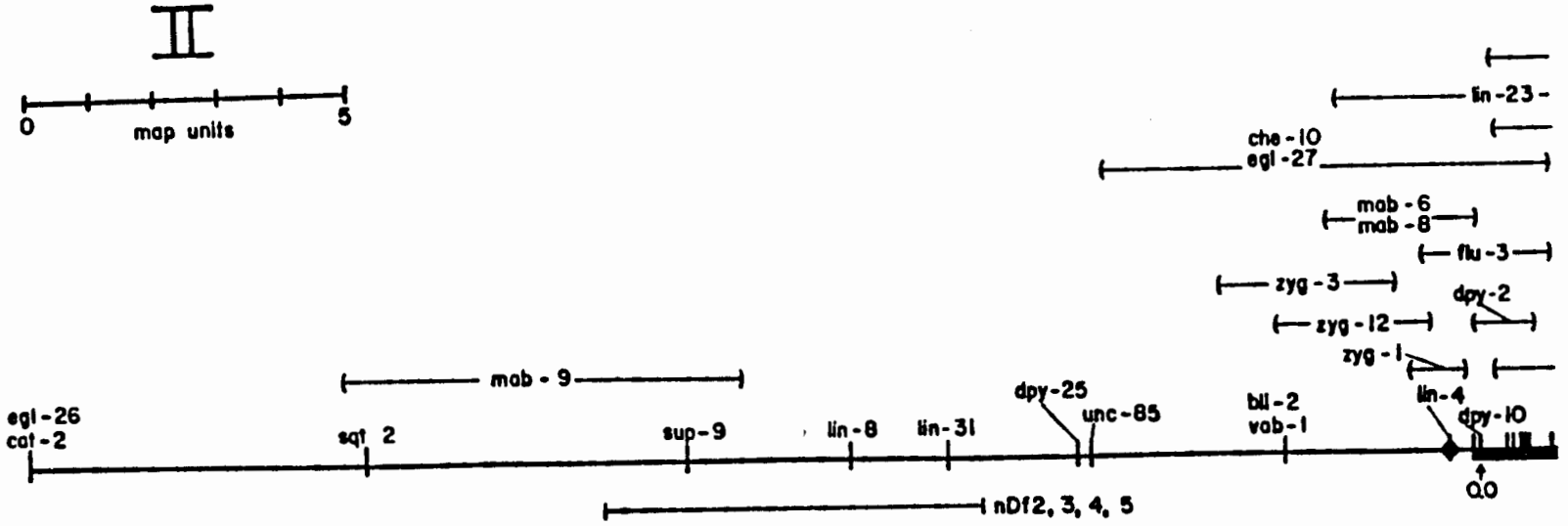
APPENDIX 1

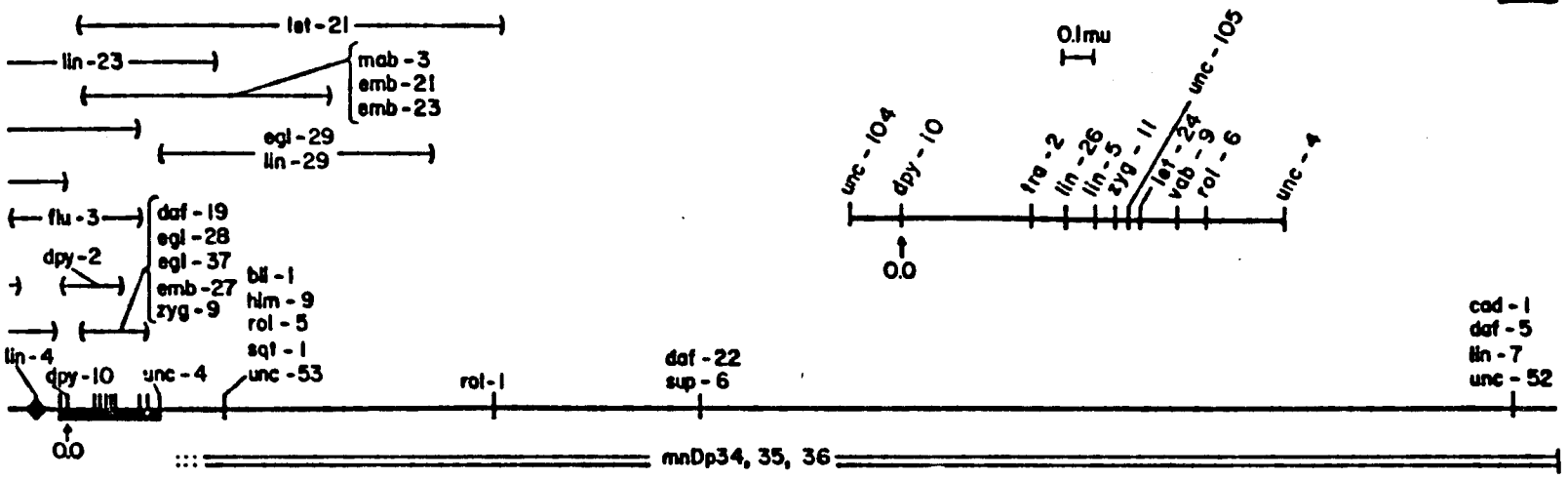
Caenorhabditis elegans

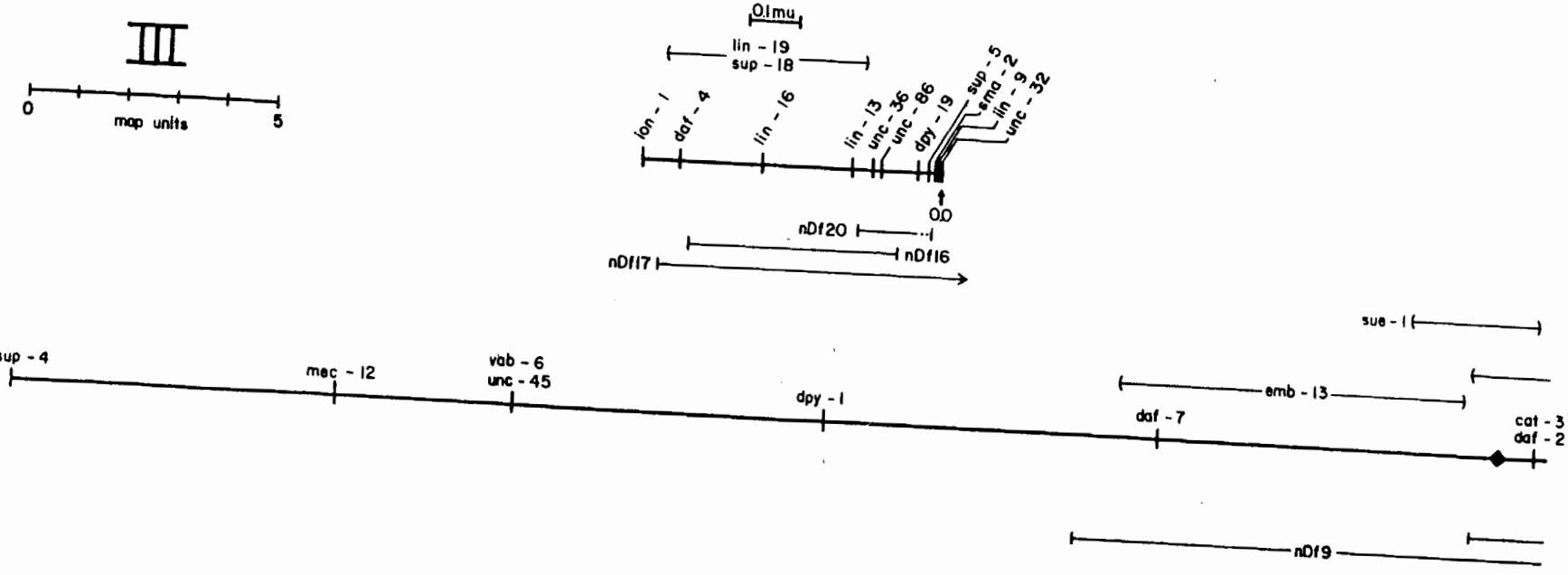
GENETIC MAP

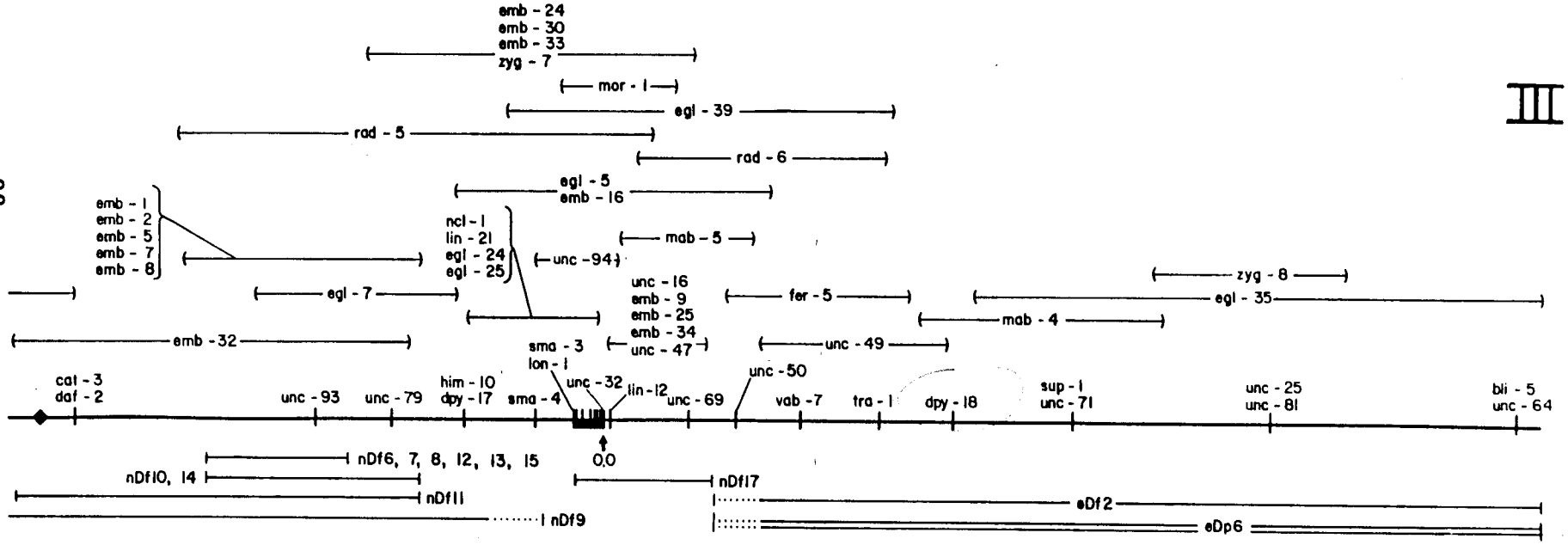

I
0
map units
5
rad-8
zyg-2
mab-2
emb-10
emb-14
emb-19
mab-1
mec-6
rad-3
unc-14
unc-55
egl-34
fer-7
emb-6
unc-89
dpy-24
unc-1
unc-63
unc-35
unc-67
fer-3
let-80
let-83
let-88
unc-57
unc-38
him-1
egl-30
unc-74
unc-13
sup-17
unc-11
unc-73
ace-2
sup-11
bli-3
lin-6
lin-17
dpy-5
unc-15
unc-56
fer-1
0.0
nDf24, 25
nDf23
sDf4
sDp2
sDp1

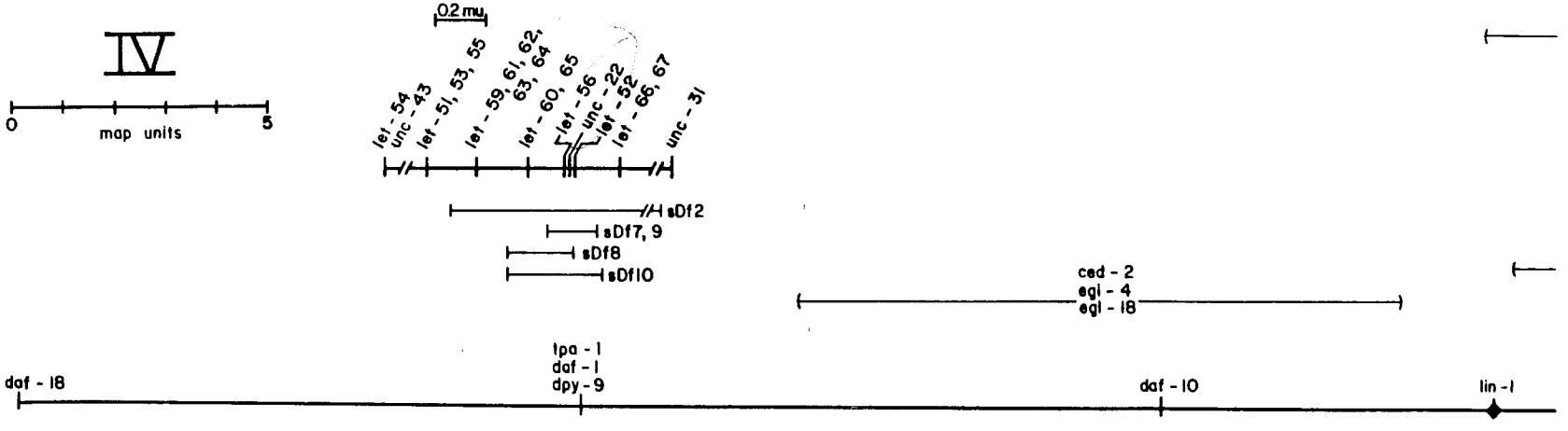

I
0.4 mu
dpy - 5
bli - 4
unc - 40
unc - 37
unc - 87
dpy - 14
unc - 15
unc - 13
let - 87
let - 84, 86
lin - 10
lin - 28
mec - 8
let - 75
unc - 56
daf - 8
0.0
sDf15
sDf16
che - 1
daf - 16
rad - 1
egl - 31
emb - 12
emb - 17
him - 2
che - 13
che - 3
fer - 6
vab - 10
dpy - 24
egl - 32
egl - 33
emb - 20
let - 80
let - 83
let - 88
daf - 17
cad - 1
let - 201
let - 202
let - 203
unc - 59
let - 204
lev - 10
let - 206
let - 207
unc - 54
let - 49
let - 50
let - 208
let - 209
let - 205
lev - 11
fer - 1
sup - 17
unc - 29
lin - 11
unc - 75
unc - 101
nDf24, 25
nDf23
eDf5
eDf10
eDf11
eDf13
eDf14
eDf15
eDf4
eDf16
eDf9
eDf17, 12
eDf6
eDf3
sDp1

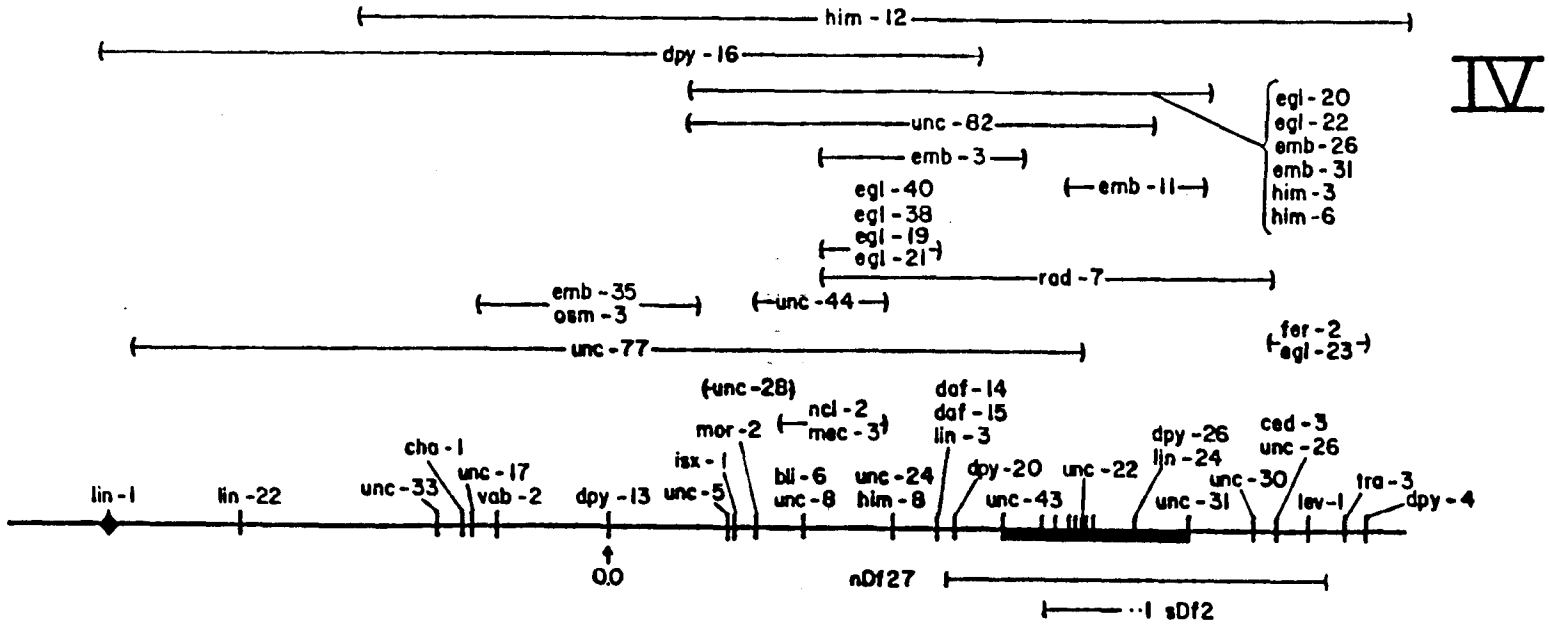

II
0
map units
5
lin-23
che-10
egl-27
mab-6
mab-8
flu-3
zyg-3
dpy-2
zyg-12
zyg-1
lin-4
dpy-10
mab-9
dpy-25
unc-85
bli-2
vab-1
egl-26
cat-2
sqt 2
sup-9
lin-8
lin-31
nDf2, 3, 4, 5
0.0

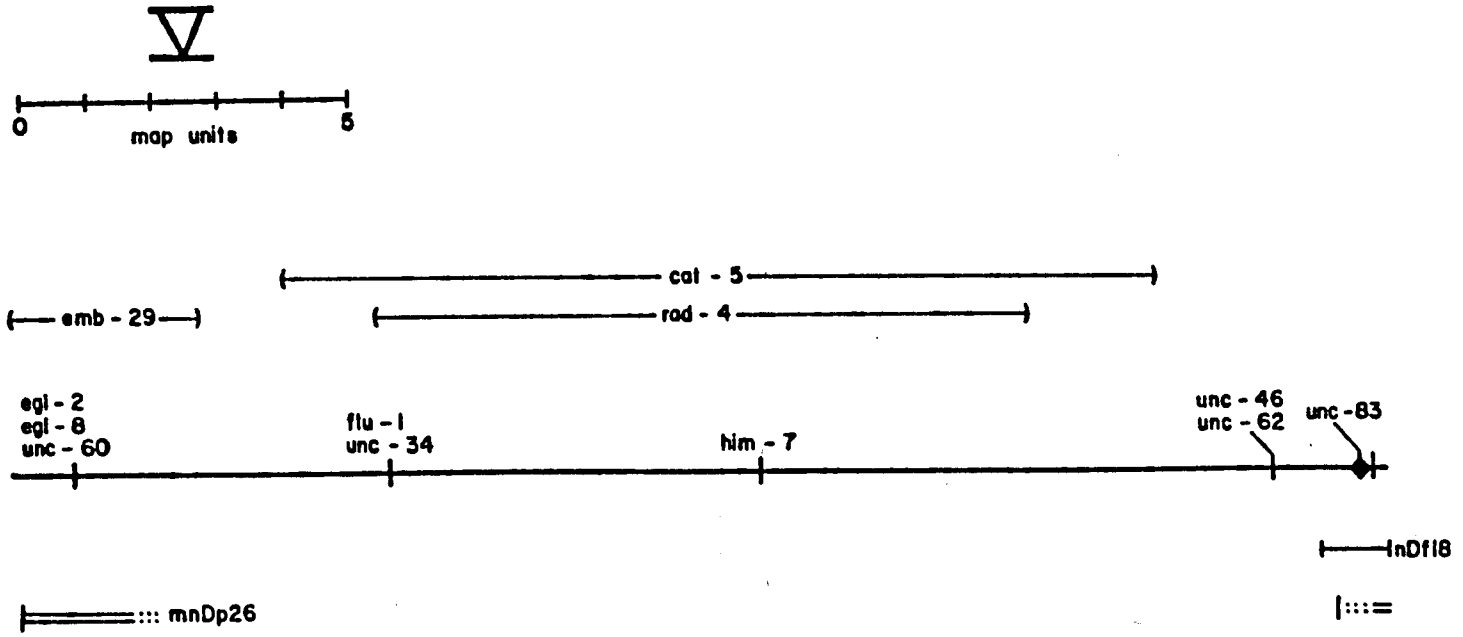

let-21
lin-23
mab-3
emb-21
emb-23
egl-29
lin-29
flu-3
daf-19
egl-28
egl-37
emb-27
zyg-9
dpy-2
bli-1
him-9
rol-5
sqt-1
unc-53
lin-4
dpy-10
unc-4
rol-1
daf-22
sup-6
cad-1
daf-5
lin-7
unc-52
0.0
mnDp34, 35, 36
0.1mu
unc-104
dpy-10
tra-2
lin-26
lin-5
zyg-11
let-24
vab-9
rol-6
unc-105
unc-4
0.0

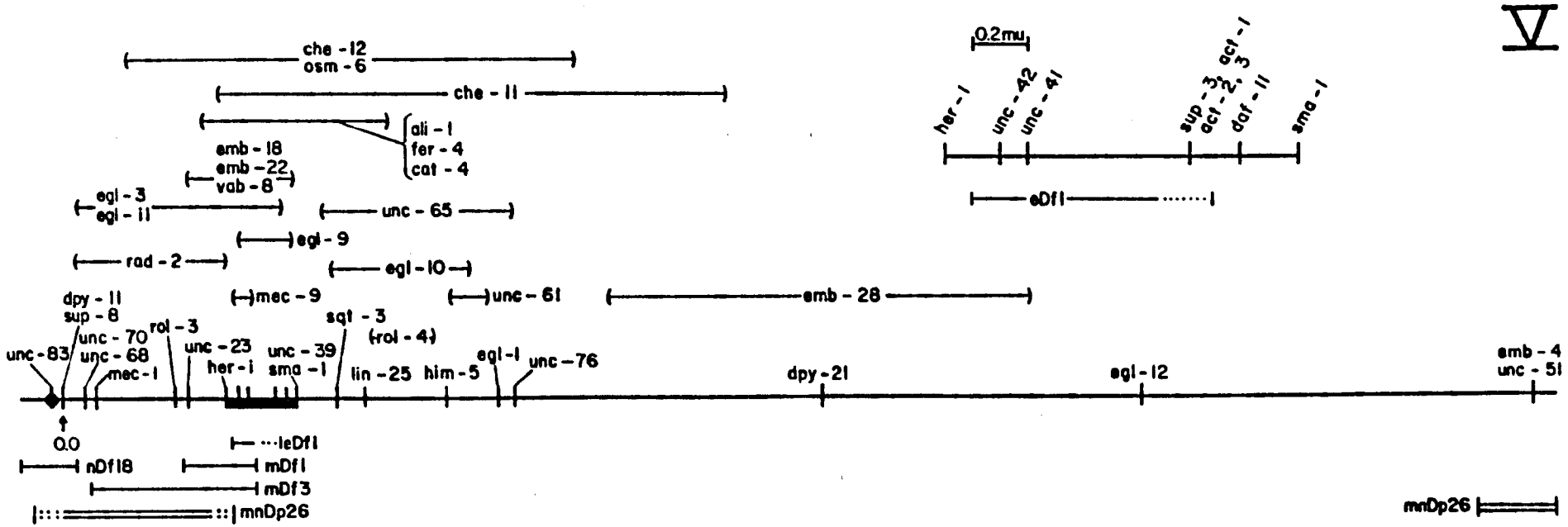

III
0
map units
5
0.1mu
lin - 19
sup - 18
lon - 1
daf - 4
lin - 16
lin - 13
unc - 36
unc - 86
dpy - 19
sup - 5
sma - 2
lin - 9
unc - 32
0.0
nDf20
nDf16
nDf17
sup - 4
mec - 12
vab - 6
unc - 45
dpy - 1
sue - 1
emb - 13
daf - 7
cat - 3
daf - 2
nDf9

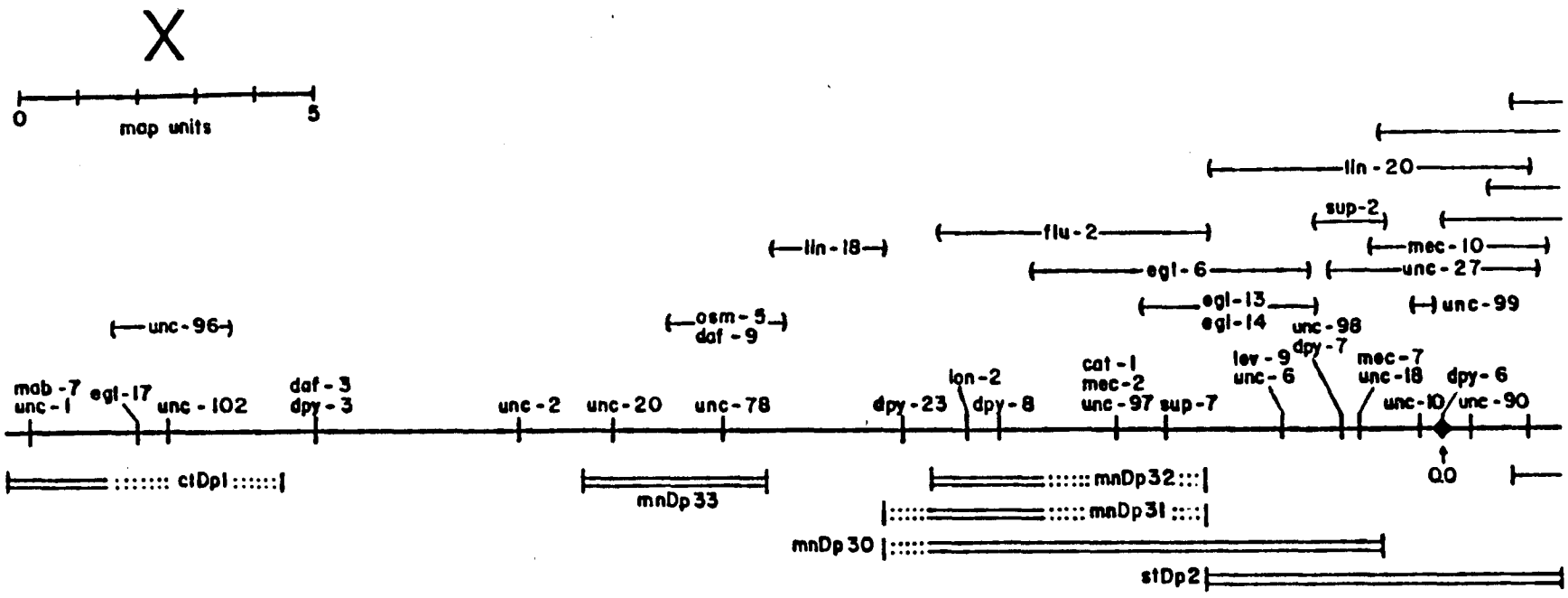

emb - 24
emb - 30
emb - 33
zyg - 7
mor - 1
egl - 39
rad - 5
rad - 6
egl - 5
emb - 16
emb - 1
emb - 2
emb - 5
emb - 7
emb - 8
ncl - 1
lin - 21
egl - 24
egl - 25
mab - 5
unc - 94
zyg - 8
egl - 7
unc - 16
emb - 9
emb - 25
emb - 34
unc - 47
fer - 5
egl - 35
mab - 4
emb - 32
sma - 3
lon - 1
unc - 49
cat - 3
daf - 2
unc - 93
unc - 79
him - 10
dpy - 17
sma - 4
unc - 32
lin - 12
unc - 69
unc - 50
vab - 7
tra - 1
dpy - 18
sup - 1
unc - 71
unc - 25
unc - 81
bli - 5
unc - 64
0.0
nDf6, 7, 8, 12, 13, 15
nDf10, 14
nDf11
nDf9
nDf17
eDf2
eDp6


IV
0
map units
5
0.2 mu
let - 54
unc - 43
let - 51, 53, 55
let - 59, 61, 62, 63, 64
let - 60, 65
let - 56
unc - 22
let - 52
let - 66, 67
unc - 31
sDf2
sDf7, 9
sDf8
sDf10
ced - 2
egl - 4
egl - 18
tpa - 1
daf - 1
dpy - 9
daf - 18
daf - 10
lin - 1

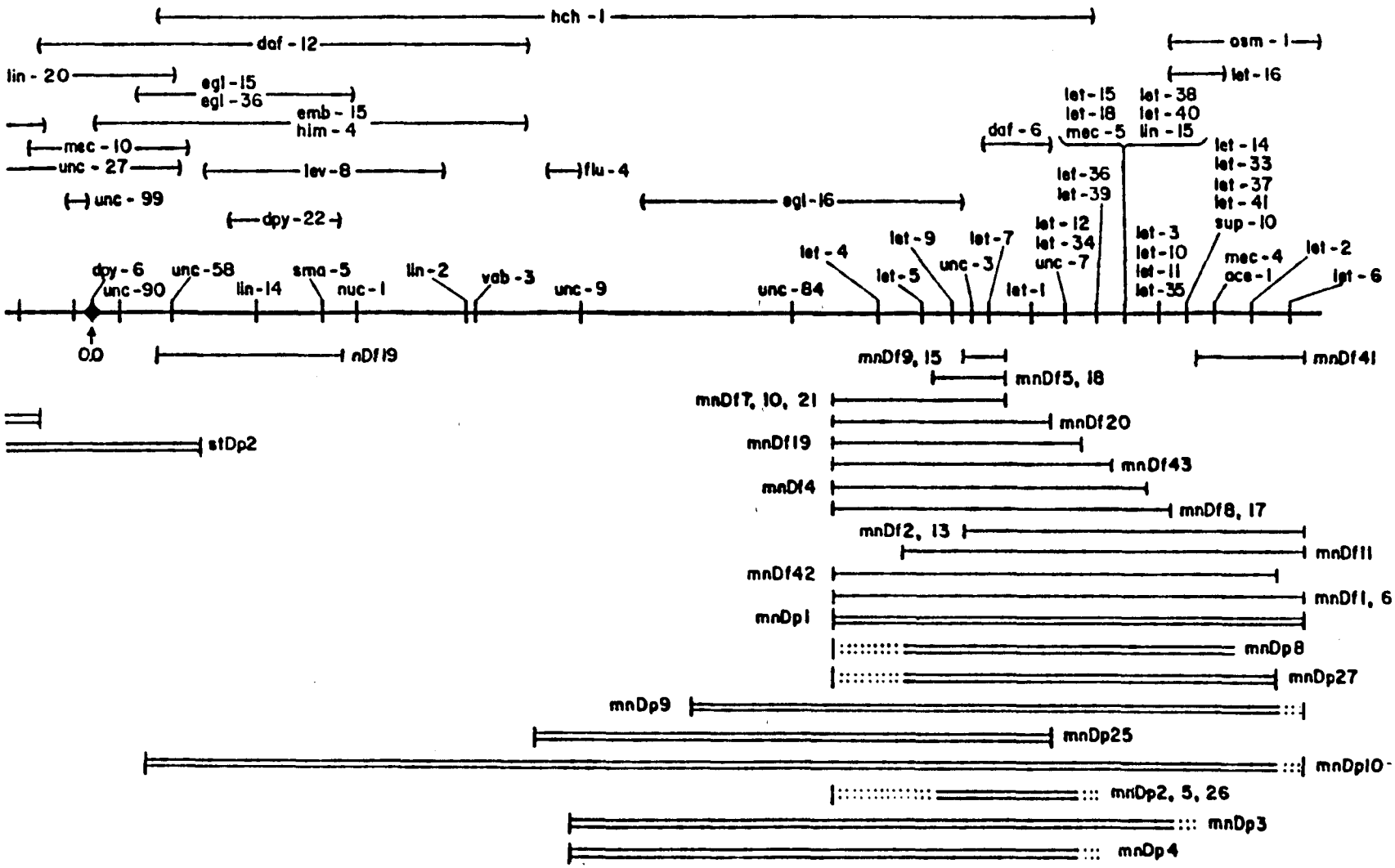

him - 12
dpy - 16
unc - 82
emb - 3
egl - 20
egl - 22
emb - 26
emb - 31
him - 3
him - 6
emb - 11
egl - 40
egl - 38
egl - 19
egl - 21
rad - 7
emb - 35
osm - 3
unc - 44
unc - 77
fer - 2
egl - 23
(unc - 28)
ncl - 2
mec - 3
daf - 14
daf - 15
lin - 3
mor - 2
cha - 1
isx - 1
unc - 5
dpy - 26
lin - 24
ced - 3
unc - 26
unc - 33
unc - 17
vab - 2
dpy - 13
bli - 6
unc - 8
unc - 24
him - 8
dpy - 20
unc - 22
unc - 43
unc - 31
unc - 30
lev - 1
tra - 3
dpy - 4
lin - 1
lin - 22
0.0
nDf27
sDf2


V
0
map units
5
cat - 5
emb - 29
rad - 4
egl - 2
egl - 8
unc - 60
flu - 1
unc - 34
him - 7
unc - 46
unc - 62
unc - 83
nDf18
mnDp26


V
0.2mu
che - 12
osm - 6
che - 11
ali - 1
fer - 4
cat - 4
emb - 18
emb - 22
vab - 8
egl - 3
egl - 11
unc - 65
egl - 9
rad - 2
egl - 10
mec - 9
unc - 61
emb - 28
dpy - 11
sup - 8
unc - 70
unc - 68
mec - 1
rol - 3
unc - 23
her - 1
unc - 39
sma - 1
sqt - 3
rol - 4
lin - 25
him - 5
egl - 1
unc - 76
dpy - 21
egl - 12
emb - 4
unc - 51
her - 1
unc - 42
unc - 41
sup - 3
act - 2, 3
act - 1
daf - 11
sma - 1
eDf1
0.0
leDf1
nDf18
mDf1
mDf3
mnDp26
mnDp26


X
0
map units
5
lin-20
sup-2
lin-18
flu-2
mec-10
egl-6
unc-27
egl-13
egl-14
unc-98
unc-99
unc-96
osm-5
daf-9
dpy-7
lev-9
unc-6
mec-7
unc-18
dpy-6
unc-10
unc-90
mab-7
unc-1
egl-17
unc-102
daf-3
dpy-3
unc-2
unc-20
unc-78
dpy-23
lon-2
dpy-8
cat-1
mec-2
unc-97
sup-7
Q0
mnDp32
mnDp31
mnDp33
mnDp30
stDp2

X


hch - 1
daf - 12
lin - 20
egl - 15
egl - 36
emb - 15
him - 4
mec - 10
unc - 27
lev - 8
unc - 99
dpy - 22
flu - 4
egl - 16
daf - 6
osm - 1
let - 16
let - 15
let - 18
mec - 5
let - 38
let - 40
lin - 15
let - 14
let - 33
let - 37
let - 41
sup - 10
let - 36
let - 39
let - 12
let - 34
unc - 7
let - 3
let - 10
let - 11
let - 35
mec - 4
oce - 1
let - 2
let - 6
let - 4
let - 9
let - 7
let - 5
unc - 3
let - 1
dpy - 6
unc - 90
unc - 58
lin - 14
sma - 5
nuc - 1
lin - 2
vab - 3
unc - 9
unc - 84
0.0
nDf19
mnDf9, 15
mnDf41
mnDf5, 18
mnDf17, 10, 21
mnDf20
stDp2
mnDf19
mnDf43
mnDf4
mnDf8, 17
mnDf2, 13
mnDf11
mnDf42
mnDf1, 6
mnDp1
mnDp8
mnDp27
mnDp9
mnDp25
mnDp10
mnDp2, 5, 26
mnDp3
mnDp4

APPENDIX 2

Derivation of the mapping equation

$$p = 1 - \sqrt{1 - [3U/(U+W)]}$$

The recombination frequency, p, is defined as the frequency of the recombinant gametes produced by an individual. In a heterozygote of the genotype, a b/+ +, the frequency of the recombinant gametes (a + and + b) is p while the frequency of the nonrecombinant gametes (a b and + +) is 1 - p. The frequencies of the recombinant and nonrecombinant progeny of an a b/+ + hermaphrodite are shown in the Punnett square below. The progeny which have a recombinant phenotype are circled.

	+ +	a b	a +	+ b
	1/2(1-p)	1/2(1-p)	1/2p	1/2p
+ + 1/2(1-p)	+ +/+ + $1/4(1-p)^2$	+ +/a b $1/4(1-p)^2$	+ +/a + 1/4p(1-p)	+ +/+ b 1/4p(1-p)
a b 1/2(1-p)	a b/+ + $1/4(1-p)^2$	a b/a b $1/4(1-p)^2$	a b/a + 1/4p(1-p)	a b/+ b 1/4p(1-p)
a + 1/2p	a +/+ + 1/4p(1-p)	a +/a b 1/4p(1-p)	a +/a + $1/4p^2$	a +/+ b $1/4p^2$
+ b 1/2p	+ b/+ + 1/4p(1-p)	+ b/a b 1/4p(1-p)	+ b/a + $1/4p^2$	+ b/+ b $1/4p^2$

The recombination frequency, p, between a and b can be calculated from the frequency of the phenotypically recombinant individuals in the progeny of an a b/+ + hermaphrodite.

$$R = \frac{\text{number of phenotypically recombinant progeny}}{\text{total progeny}}$$

R = frequency of recombinant progeny

$$R = 1/4p(1-p) + 1/4p(1-p) + 1/4p(1-p) + 1/4p(1-p) + 1/4p^2 + 1/4p^2$$

$R = p(1-p) + 1/2p^2$

$R = p - 1/2p^2$

$0.5p^2 - p + R = 0$

solve for p using: $-b - \sqrt{-b^2 - 4ac} \;/\; 2a$

$p = 1 - \sqrt{1 - 2R}$

In the case of a let-a(sx)unc-22(s7)/+ $\pm$ hermaphrodite where 1/4 of the progeny, including half of the recombinants, die the total number of progeny is calculated as 4/3(Wild + Unc-22) and the number of recombinant progeny as 2(Unc-22).

Therefore:

$R = 2(\text{Unc-22}) / [4/3(\text{Wild} + \text{Unc-22})]$

$R = 3U/2(U + W)$

and

$p = 1 - \sqrt{1 - 2[3U/2(U + W)]}$

$p = 1 - \sqrt{1 - [3U/(U + W)]}$

APPENDIX 3

Estimation of the total number of essential genes uncovered by sDf2

The estimate of the total number of genes in the region uncovered by sDf2 was determined using the method employed by Meneely and Herman (1979; 1981). This method assumes that the number of mutations per essential gene in this region follows a Poisson distribution.

The total number of genes was estimated using the fraction of mutated genes represented by more than one mutant (f). According to the Poisson distribution

$$f = (1 - e^{-m} - me^{-m})/(1 - e^{-m})$$

where m is the average number of mutations per essential gene and e^{-m} is the fraction of genes not mutated.

In this study, twenty essential genes were identified, eight of which are represented by more than one mutation. For $f = 8/20$; $m = 0.95$ and $e^{-m} = .39$ thus, the twenty essential genes identified represent 61% of the total number of genes in this region. The estimate of the total number of genes in this region is 32.

APPENDIX 4

List of strains carrying the lethal and sterile mutations used in this thesis

Strain	BC #	Gene(mutation)
S-H89*S1	0987	let-52(s42)
S-H90*S1	0988	let-56(s46)
S-H91	0330	let-51(s41)
S-H93*S1	1223	let-57(s47)
S-H94*S1	0961	let-59(s49)
S-H95*S1	0916	let-60(s59)
S-H96*S1	0917	let-61(s65)
S-H100*S1	1167	let-72(s52)
S-H102*S1	1224	let-56(s50)
S-H103*S1	1168	s51
S-H104*S1	0905	let-64(s171)
S-H105*S1	0918	let-63(s170)
S-H115*S1	0919	let-65(s174)
S-H116	0424	s169
S-H117*S1	0906	let-56(s168)
S-H118*S1	0932	let-59(s172)
S-H119*S1	0933	let-66(s176)
S-H120*S1	0934	let-62(s175)

S-H122*S1	0989	s166
S-H123*S1	0991	s212
S-H124*S1	0990	s213
S-H125*S1	0953	let-67(s214)
S-H126*S1	0992	s215
S-H127*S1	0954	let-64(s216)
S-H203*S1	0962	let-65(s254)
S-H206*S1	0963	let-56(s173)
S-H233*S1	0904	s63
S-H374*S1	1110	let-92(s504)
S-H530	1098	s673
S-H531	1099	s674
S-H532	1100	s675
S-H533	1101	s676
S-H534	1102	let-92(s677)
S-H535	1103	let-91(s678)
S-H536	1104	let-63(s679)
S-H537	1105	let-68(s680)
S-H538	1106	s681
S-H539	1107	s682
S-H540	1108	s683
S-H541	1109	let-69(s684)
S-H542	1111	let-73(s685)
S-H543	1112	s686
S-H545	1114	s688
S-H546	1115	let-70(s689)

S-H547	1116	s690
S-H548	1118	s691
S-H549	1119	let-71(s692)
S-H550	1120	let-68(s693)
S-H551	1121	let-65(s694)
S-H552	1122	let-72(s695)
S-H553	1123	let-68(s696)
S-H554	1124	let-74(s697)
S-H555	1125	s698

BIBLIOGRAPHY

Alberts, B., D.Bray, J.Lewis, M.Raff, K.Roberts and J.D.Watson. 1983.Molecular Biology of the Cell. Garland Publishing, Inc. New York.

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

Belyaeva, E.S., M.G.Aizenzon, V.F.Semeshin, I.I.Kiss, K. Koczka, E.M.Baritcheva, T.D.Gorelova and I.F.Zhimulev. 1980. Cytogenetic analysis of the 2B3-4 - 2B11 region of the X chromosome of Drosophila melanogaster. Chromosoma 81:281-306.

Bingham, P.M., R.Levis and G.M.Rubin. 1981. The cloning of DNA sequences from the white locus of Drosophila melanogaster using a novel and general method. Cell 25:693-704.

Chambon, P. 1981. Split genes. Sci. Amer. 224:60-71.

Chovnick, A., G.H.Ballantyne and D.G.Holm. 1971. Studies on gene conversion and its relationship to linked exchange in Drosophila melanogaster. Genetics 69:179-209.

Chovnick, A., W.Gelbart and M.McCarron. 1977. Organization of the rosy locus in Drosophila melanogaster . Cell 11:1-10.

Coulondre, C. and J.H.Miller. 1977. Genetic studies of the lac repressor. IV. Mutagenic specificity in the lacI gene of Escherichia coli. J.Mol.Biol. 117:557.

Derman, E., K.Krauter, L.Walling, C.Weiberger, M.Ray and J.E.Darnell, Jr. 1981. Transcriptional control in the production of liver-specific mRNAs. Cell 23:731-739.

Early, P., J.Rogers, M.Davis, K.Calame, M.Bond, R.Wall and L.Hood. 1980. Two mRNAs can be produced from a single immunoglobulin u gene by alternative RNA processing pathways. Cell 20:313-319.

Fink, G.R. and G.A.Styles. 1974. Gene conversion of deletions in the HIS 4 region of yeast. Genetics 77:231-244.

Finnerty, V. 1976. Genetic units of Drosophila - simple cistrons. pp 721-760. In: The genetics and biology of Drosophila, Vol 1b. Edited by M.Ashburner and E.Novitski. Academic Press, London.

Garrells, J.I. 1979. Changes in protein synthesis during myogenesis in a clonal cell line. Dev.Biol.73:134-152.

Gausz, J., G.Bencze, H.Gyurkovics, M.Ashburner, D.Ish-Horowicz and J.J.Holden. 1979. Genetic characterization of the 87C region of the third chromosome of Drosophila melanogaster. Genetics 93:917-934.

Gross, P.R. 1967. Curr.Topics Dev.Biol.2:1

Gurdon, J.B. 1962. The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. J. Embryol. Exp. Morphol. 10:622-640.

Herman, R.K. 1978. Crossover suppressors and balanced recessive lethals in Caenorhabditis elegans. Genetics 88:49-65.

Herman, R.K., D.G.Albertson and S.Brenner. 1976. Chromosome rearrangements in Caenorhabditis elegans. Genetics 83:91-105.

Herman, R.K., J.E.Madl and C.K.Kari. 1979. Duplications in Caenorhabditis elegans. Genetics 92:419-435.

Herman, R.K. and H.R.Horvitz. 1980. Genetic analysis of Caenorhabditis. pp 227-261 In: Nematodes as Biological Models, Vol. 1 Edited by B.M.Zuckerman. Academic Press, New York.

Herskowitz, I.and D.Hagen. 1980. The lysis-lysogeny decision of phage lambda: explicit programming and responsiveness. Annu. Rev. Genet. 14:399-445.

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

Hochman, B. 1978. The fourth chromosome of Drosophila melanogaster. In: The genetics and Biology of Drosophila , Vol. 1b Edited by M.Ashburner and E.Novitski. Academic Press, London.

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

Isenberg, I. 1979. Histones. Annu. Rev. Biochem. 48:159-191.

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

Karn, J., S.Brenner and L.Barnett. 1983. Protein structural domains in the Caenorhabditis elegans unc-54 myosin heavy chain gene are not separated by introns. P.N.A.S.

80:4253-4257.

Kramer, J.M., G.N.Cox and D.Hirsh. 1982. Comparison of the complete sequences of two collagen genes from Caenorhabditis elegans. Cell 30:599-606.

Lewis, J.A., C.H.Wu, J.H.Levine and H.Berg. 1980. Levamisole-resistant mutants of the nematode Caenorhabditis elegans appear to lack pharmacological acetycholine receptors. Neuroscience 5:967-989.

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. J.Mol.Biol.114:133-140.

Macleod, A.R., J.Karn and S.Brenner. 1981. Molecular analysis of the unc-54 myosin heavy chain gene of Caenorhabditis elegans. Nature 290:386-390.

McCarron, M., J.O'Donnel, A.Chovnick, B.S.Bhullar, J.Hewitt and E.P.M.Candido. 1979. Organization of the rosy locus in Drosophila melanogaster: further evidence in support of a cisacting control element adjacent to the xanthine dehydrogenase structural element. Genetics 91:275-293.

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 Caenorhabditis elegans. Genetics 97:65-84.

Miller, J.H. and W.S.Reznikoff eds. 1978. The Operon. New York: Cold Spring Harbor Laboratory.

Moerman, D.G. 1980. A genetic analysis of the unc-22 region in Caenorhabditis elegans. Ph.D. thesis, Simon Fraser University, Burnaby, British Columbia.

Moerman, D.G. and D.L.Baillie. 1979. Genetic organization in Caenorhabditis elegans: fine-structure analysis of the unc-22 gene. Genetics 91:95-103.

Moerman D.G. and D.L.Baillie. 1981. Formaldehyde mutagenesis in the nematode Caenorhabditis elegans. Mutat.Res.80:273-279.

Moerman, D.G., S.Plurad, R.H.Waterston and D.L.Baillie. 1982. Mutations in the unc-54 myosin heavy chain gene of Caenorhabditis elegans that alter contractility but not muscle structure. Cell 29:773-781.

O'Donnell, J., H.C.Mandel, M.Kraus and W.Sofer. 1977. Genetic

and cytogenetic analysis of the adh region in Drosophila melanogaster. Genetics 86:553-566.

Perry, R.P. 1981. RNA processing comes of age. J.Cell.Biol. 91:28s-38s.

Ptashne, M., A.Jeffrey, A.D.Johnson, R.Maurer, B.J.Meyer, C.O.Pabo, T.M.Roberts and R.T.Sauer. 1980. How the lambda repressor and cro work. Cell 19:1-11.

Rasch, E.M., H.J.Barr and R.W.Rasch. 1971. The DNA content of sperm of Drosophila melanogaster. Chromosoma 33:1-18.

Riddle, D.L. 1978. The genetics of development and behavior in Caenorhabditis elegans. J.Nematol.10:1-16.

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

Rose, A.M. 1980. Genetic studies on the gene coding for paramyosin in Caenorhabditis elegans: unc-15 and the adjacent region. Ph.D. thesis, Simon Fraser University, Burnaby, British Columbia.

Rose, A.M. and D.L.Baillie. 1979a. The effect of temperature and parental age on recombination and nondisjunction in Caenorhabditis elegans. Genetics 92:409-418.

Rose, A.M. and D.L.Baillie. 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., D.L.Baillie, E.P.M.Candido, K.A.Beckenbach and D.Nelson. 1982. The linkage mapping of cloned restriction fragment length differences in Caenorhabditis elegans. Mol.Gen.Genet.188:286-291.

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). Mutat.Res.110:39-48.

Royal, A., A.Garapin, B.Cami, F.Perrin, J.L.Mandel, M.LeMeur, F.Bregere, J.P.LePennec, P.Chambon and P.Kourilsky. 1979.

The ovalbumin gene region: common features in the organization of three genes expressed in chicken oviduct under hormonal control. Nature 279:125-132.

Spradling, A.C. and G.M.Rubin. 1981. Drosophila genome organization: conserved and dynamic aspects. Ann.Rev.Genet.15:219-264.

Snutch, T.P. and D.L.Baillie. 1983. Alterations in the pattern of gene expression following heat shock in the nematode Caenorhabditis elegans . Can.J.Biochem.Cell Biol.61:480-487.

Strickberger, M.W. 1968. Genetics. The Macmillan Company, New York.

Sulston, J.E. and S.Brenner. 1974. The DNA of Caenorhabditis elegans. Genetics 77:95-104.

Sulston, J.E. and H.R.Horvitz. 1977. Post-embryonic cell lineages of the nematode Caenorhabditis elegans. Develop.Biol.56:110-156.

Tsai, S.Y., M-J.Tsai, C-T.Lin and B.W.O'Malley. 1979. Effect of estrogen on ovalbumin gene expression in differentiated nontarget tissues. Biochemistry 18:5726-5731.

von Ehrenstein, G. and E.Schrierenburg. 1980. Cell lineages and development of Caenorhabditis elegans and other nematodes. pp 1-72. In: Nematodes as Biological Models. Vol 1. B.M.Zuckerman (ed), Academic Press, New York.

Waterston, R.H. and S.Brenner. 1978. A suppressor mutation in the nematode Caenorhabditis elegans acting on specific alleles of many genes. Nature 275:715-719.

Waterston, R.H., R.M.Fishpool and S.Brenner. 1977. Mutants affecting paramyosin in Caenorhabditis elegans. J.Mol.Biol. 117:679-698.

Waterston, R.H., J.N.Thomson and S.Brenner. 1980. Mutants with altered muscle structure in Caenorhabditis elegans Develop.Biol. 77:271-302.

Waterston R.H. 1981. A second informational suppressor, sup-7 X , in Caenorhabditis elegans. Genetics 97:307-325.

Waterston, R.H., K.C.Smith and D.G.Moerman. 1982. Genetic finestructure analysis of the myosin heavy chain gene, unc-54, of Caenorhabditis elegans. J.Mol.Biol.158:1-15.

Woodruff, R.C. and M.Ashburner. 1979. The genetics of a small autosomal region of Drosophila melanogaster containing the

structural gene for alcohol dehydrogenase. II. Lethal mutations in the region. Genetics 92:133-149.

Wright, T.R.F., W.Beermann, J.L.Marsh, C.P.Bishop, R.Steward, B.C.Black, A.D.Tonsett and E.Y.Wright. 1981. The genetics of dopa decarboxylase in Drosophila melanogaster. IV. The genetics and cytology of the 37B10 - 37D1 region. Chromosoma 83:45-58.

Young, M.W. and B.H.Judd. 1978. Nonessential sequences, genes and the polytene chromosome bands of Drosophila melanogaster. Genetics 88:723-742.

Zengel, J.M. and H.F.Epstein. 1980a. Muscle development in Caenorhabditis elegans: A molecular genetic approach. pp 73-126 In: Nematodes as Biological Models. Vol 1 B.M.Zuckerman (ed). Academic Press, New York.

Zengel J.M. and H.F.Epstein. 1980b. Identification of genetic elements associated with muscle structure in the nematode Caenorhabditis elegans. Cell Motility 1:73-97.

Zhimulev, I.F., G.V.Pokholkova, A.V.Bgatov, V.F.Semeshin and E.S. Belyaeva. 1981. Fine cytogenetic analysis of the band 10A1-2 and the adjoining regions in the Drosophila melanogaster chromosome. II. Genetical analysis. Chromosoma 82:25-40.