GENETIC VARIATION AND POPULATION STRUCTURE IN WILD ISOLATES OF *Caenorhabditis elegans* COLLECTED FROM CALIFORNIA

by

ALIVIA DEY

A thesis submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

and

The Graduate School of Biomedical Sciences

University of Medicine and Dentistry of New Jersey

in partial fulfillment of the requirements

for the degree of

Master of Science

Graduate Program in Microbiology and Molecular Genetics

written under the direction of

Professor Jody Hey

and approved by

_____

_____

_____

New Brunswick, New Jersey

[October, 2007]

ABSTRACT OF THE THESIS

Genetic variation and population structure in wild isolates of *Caenorhabditis elegans* collected from California

By ALIVIA DEY

Thesis Director: Prof. Jody Hey

Though much widespread in distribution, the nematode worm C*aenorhabditis elegans* exhibits low levels of genetic variation at the DNA sequence level, a paradox which may be partly explained by its typical self-fertilization mode of reproduction. However recent work on the genetic structure of natural populations of *C. elegans* from France, Scotland and Portugal suggests a substantial level of outcrossing together with finding of a very strong local population structure. To find out whether this same occurrence extrapolates to other wild isolates of worm populations, the present study was carried on worm populations from California. Here, genetic variation is studied by detecting single nucleotide polymorphisms in a random genome-wide manner using Amplified Fragment Length Polymorphism analysis (AFLP). The worms used in this study are natural isolates of *C. elegans* collected from parks and gardens around the Los Angeles area in southern California. Some populations sampled were a few meters apart, enabling the assessment of variation and population structure on a very local scale. As previous studies, a low overall genetic diversity was found with these worm populations. The finding of a strong population structure and high linkage disequilibrium both within and between chromosomes was also concurrent with the findings from other wild isolates. All these findings elucidate the extra-ordinary evolutionary dynamics of these sets of worm from the wild.

Acknowledgements

Like any other wary student who started out in graduate school & has often at times become overwhelmed by it, I too wasn't an exception. In retrospect today, as I reflect on these couple years, I cannot imagine myself but continuing forward to be in the challenge of it, doing research in the subject that excites me most. For this transition of perceptions, I am deeply indebted to my wonderful advisor, mentor & teacher– Prof. Jody Hey for his incredible inspiration, encouragement, guidance and support through the couple years I have worked in his lab. He has in every way been there, be it scholarly, emotionally, socially, and administratively. I appreciate the endless time, patience, advice & useful suggestions he had shared with me during the course of this research work & keeping my morale high even at times of occasional frustrations of it. I thank him for giving me a strong feel for evolutionary biology and population genetics as a subject, thus opening up a whole new world for me. With his enthusiasm & sincere efforts in explaining things clearly, simply to me & taking care that I learn, I have indeed learnt more through my discussions with him than through my course-work or books. I can not overstate the importance of him in my graduate career & more importantly in being a great role model in life.

I wish to thank my committee members – Drs Chi-hua Chiu and Dr. Andy Singson for agreeing to be in my committee at a short notice. I am in many more ways thankful than I can relate, to Dr. Chi-hua Chiu for her enthusiastic, academic & emotional support and her invaluable friendship right from the first day in Rutgers.

It would seem ungrateful on my part not to thank the attendees of the Evolution Discussion Group meetings through the years. I owe much to this group for giving me a

feel for the science one loves to pursue as well as be a critic of it. Thanks particularly to Prof. Hey & Prof. Robert Trivers for those long sessions of theory, background related to the literature and the priceless fun stories behind seminal works.

I would like to thank my labmate Yong Wang for providing me such a fun and stimulating environment to work in and for his invaluable friendship. I'll greatly miss his company in sharing my lab adventures, discussions, debates & our office space together.

Thanks are due to the wonderful faculty and very adept administrative staff of the Graduate Program in Molecular Biosciences for their much appreciated help, and more so to an international student - for making it feel like home.

I must thank Antoine Barriere at the Institut Jaques Monod, France whose help I am extremely grateful for during the course of working with the experiments. I have much benefited from the series of protocol sharing & comparing that we had over emails. I doubt this work would have been finally realised had he not been there.

I thank Arjun Sivasundar for getting me started with the worm work and to the undergraduate students Nikul Patel and Melissa Wang who assisted me during the course of it.

I am much indebted to my friends at Rutgers University, UMDNJ and Asha for Education, CNJ who have made graduate life all the more fun & memorable. These individuals always helped me to keep my life in context & helped me sail gracefully through some very tough times. I thank them for their lifelong friendship.

Finally, to my loving mom – thank you for just being there, encouraging me through & for keeping it up boldly. I love you so much.

- *Alivia Dey, July 2007*

Table of Contents

Lists of tables

List of illustrations

Chapter 1

Introduction:

With rapid advances in genomics in the beginning of the twenty-first century, a changed view of the biological world is apparent which has ushered in new links between molecular biology, ecology and evolution. Genomic studies of the model organisms have been a crucial ingredient in this synthesis allowing studies of natural variation acting as a bridge between molecular analyses of gene function and evolutionary investigations of adaptation and natural selection. The nematode worm, *Caenorhabditis elegans* has become one such key player in model-systems research and has contributed vastly to our understanding of fundamental mechanisms in areas of animal development, genetics, neurobiology, developmental biology, behaviour, genome architecture and function. However, in contrast to other model systems like *Drosophila* where the history of evolutionary genetics can be traced back to the 30s, and *Arabidopsis* in the 1960s, little progress have been made to study its ecology, natural history, evolution and natural variation of this species. The scenario is slowly beginning to change with more researchers beginning to develop an understanding of the evolutionary mechanism in *C. elegans* (Haag et al. 2007; Cutter 2006; Sivasundar and Hey 2005; Barriere and Felix 2005; Haber et al. 2005; Denver et al. 2004; Kiontke et al. 2004; Graustein et al 2002; Hodgkin and Doniach 1997).

This soil nematode (Nematoda: Rhabditidae: Peloderinae) occurs in soil almost worldwide (Abdul Kader and Cote, 1996; Hodgkin and Doniach, 1997) and is one of the nearly 20 species of bactivorous nematodes in the genus *Caenorhabditis* (Sudhaus and Kiontke 1996). It is a free living nematode, which has been isolated from garden soils as

well as from compost piles with only recently being also isolated from rotting fruits, and is found to be associated with other soil invertebrates such as snails (Barriere and Felix 2007). *Caenorhabditis elegans* and three other species of this genus, *C briggsae*, *C. remanei* and *C. sp*. (strain CB5161) form a distinct clade within the family Rhabditidae (Fitch et al.1995). The phylogenetic relationships between these closely related species have only recently been resolved, with *C. elegans* being basal to the other three species in this clade (Fig 1.1).

Although *C. elegans* and its closest relatives *C. briggsae* and *C. remanei* are morphologically very similar yet they are vastly different at the DNA sequence level. Though belonging to the same genus & both being androdioceous, the most recent common ancestor of *C. briggsae* & *C. elegans* existed 80–110 million years ago (mya), and thus they are more evolutionarily distant than, for example, human and mouse (divergent time for the latter being about 75-85 mya (Hillier et al. 2007; Nei et al., 2001). This lack of closely related species to *C.* elegans confounds phylogenetics as well as poses hindrance to comparative evolutionary studies.

Besides, an interesting feature for evolutionary biology is the peculiar mode of reproduction of *C .elegans.* The androdioecious species, *C. elegans* & *C. briggsae*, reproduce largely by self-fertilisation of hermaphrodites, whereas in contrast *C. remanei* exhibits an obligate gonochoristic breeding system. Populations of *C. elegans* consist of two sexes, selfing hermaphrodites (XX) and facultative males (XO) which are able to cross with the hermaphrodites. Males arise rarely, due to non-disjunction of the X-chromosome during meiosis (Meyer 1997). Thus when males do arise, they could

potentially persist in the population. However laboratory observations show that mating efficiency is poor in this species compared to that of the gonochristic species. In fact, *C. elegans* males mate more efficiently with females of another species, *C. remanei*, than they do with conspecific hermaphrodites (Chasnov and Chow 2002), despite the fact that no viable progeny result (Baird et al., 1992; Baird et al., 1994). When males do arise in laboratory populations, they are selected against and eliminated from the population at an approximately exponential rate (Stewart and Philips 2002). Consequently in laboratory reproduction in worm population is almost entirely by selfing. This unique mating system and its relationship to genetic variation raises relevant questions regarding population studies in *C. elegans* and evolution of sex, a much enigma in evolutionary biology which altogether makes it an excellent system to study.

Although evolutionary and ecological studies of *C. elegans* have been scarce until a few years ago, new studies over the recent years have made an important contribution to our knowledge of molecular polymorphisms, genetic diversity, genetic structure of populations, demographics and reproduction in natural populations of the worms. This thesis is an attempt to summarise what is currently known in the field of *C. elegans* population genetics, with identification of key gaps in our knowledge of their biology that forms the background for this research. To address these gaps, the research presented here augments some of the recent observations made and aims to provide baseline data on the genetic variation in *C. elegans*, to examine the population structure and the pattern of linkage disequilibrium in natural isolates of the worms collected from California.

Background & Significance

1.1 Molecular Polymorphisms and Genetic variation of *C. elegans*

Previous work carried on molecular polymorphisms in *C. elegans* was based on a set of natural isolates that were collected worldwide over many years, each kept as a single isogenic strain available at Caenorhabditis Genetics Center (CGC). Collected from North America (mostly California), Western Europe (including N2, the reference strain isolated in Bristol, U.K) and Adelaide (Australia), these isolates were used as representatives to assess molecular and phenotypic variation within species, even though a significant part of the species is missing due to restrictive geographic sampling. *Caenorhabditis elegans* has never been isolated in Asian samples (unlike *C. briggsae*) and not much sampling was attempted in South America.

Recently, however sets of wild isolates have been collected from parts of Germany (Haber et al., 2005), France (Barrière and Félix, 2005), Scotland (A Cutter, 2005), California (Sivasundar and Hey, 2005) and parts of Africa (Elie Dolgin, M.A.Felix and A.D. Cutter., unpublished data). Several strains and lines were established from each collection yielding appreciable information on the diversity within a local population.

a. Low Global Genetic Diversity of C. elegans

Genetic polymorphism have been found to be low in *C. elegans* both at the mitochondrial sequence level (Denver et al., 2003), some nuclear genes (Graustein et al., 2002; Denver et al., 2003), at microsatellite markers (Sivasundar & Hey, 2003; Haber et

al., 2005, Barriere and Felix 2005), from shotgun sequencing data (Swan et al., 2002; Wicks et al., 2001) and SNP data (Koch et al. 2000; Cutter 2006). At nuclear loci, the average differences between gene copies has been found to be close to .001/bp which is typically similar to humans and about 20 fold lower than *Drosophila melanogaster* (Cutter 2006; Denver et al., 2003). Diversity at the whole genome level was measured by AFLP (Amplified Fragment length Polymorphism) on natural populations isolated from France. The average nucleotide diversity (π) was found to be .81 x 10^{-3} (Barriere and Felix 2005).

Contrastingly, the mutation rate both at the mitochondrial and nuclear level in the mutation accumulation (MA) line experiments have not been found to be that low. Typically it is of the order of 9.7 X 10^{-8} mutation per site per generation for the mitochondrial loci and 2.1 X 10^{-8} for the nuclear loci (Denver et al. 2000; Denver et al. 2004). But this can be accounted for the fact that MA lines represent raw mutational process, whereas diversity among the wild isolates is the result of mutation as well as population structure and phenotypic selection.

At the chromosomal level, polymorphisms are more frequent in distal arm regions than in the central clusters (Koch et al., 2000) interpreting this as evidence that genes on the autosomal arms experience more rapid evolution.

b. Effective population size (N_e)

An important parameter in population genetics is the effective population size (N_e), which represents the number of individuals of an equivalent ideal random-mating population under mutation and genetic drift, or the number of individuals contributing to

the next generation. It can be estimated using the mutation rate μ and the observed molecular diversity π from $\pi = 4N_e\mu$. For a given mutation rate, it is thus equivalent to a measure of molecular diversity. Practically, N_e is reduced by inbreeding or non-random mating due to geographical structure, and is almost always smaller than the number of breeding individuals (Hartl and Clark, 1997; Charlesworth et al., 2003). The low genetic variation in *C. elegans* suggests a small effective population size. Available estimates of N_e from microsatellites for the whole species range from 200 to 44,000, depending on which locus is considered (Sivasundar and Hey, 2003). From AFLP data, it was estimated to be 9,600 for the four French samples (Barrière and Félix, 2005). These numbers are surprisingly small for such a widespread species with large local populations especially considering that these worms reach densities of 10 individuals / g in a compost heap; (Barrière and Félix, 2005) and on decaying mushroom compost numbers are far larger (Cutter, 2006). Estimating N_e from π derived from CGC strains and a single individual from each of the European population yields N_e= 80,000. These values are substantially higher than N_e inferred from microsatellites and AFLP data, but are still rather small relative to the high census densities that these nematodes can achieve.

Thus multiple lines of evidence, from different classes of molecular markers, now point to a pattern of *C. elegans* low genetic diversity, almost comparable to that of humans and 20-fold lower than that found in *Drosophila melanogaster*. In general, widespread species with large population sizes are expected to harbour large amounts of genetic variation; however in organisms such as *C. elegans,* which reproduce primarily by selfing, variation within a population may be reduced and the effective population size too maybe reduced by up to a factor of two (Pollak, 1987). A low mutation rate cannot

be invoked as an explanation for such a reduced variation, given the high mutation rates estimated for mutation accumulation lines (Denver et al. 2000; Denver et al. 2004).The level of genetic diversity also depends on the structure of natural populations. For example, *Homo sapiens* have low nucleotide diversity, due to a recent bottleneck in the population. But the pattern of microsatellite diversity of *C. elegans* does not show evidence for a recent global bottleneck followed by population expansion. In the light of selection, such low variation can also be stemmed from either a selective sweep or a background selection at linked loci. However, a selective sweep would cause a similar effect on patterns of variation as would a population expansion. The fact that none of the analyses found evidence for population growth argues against such a scenario (Sivasundar and Hey, 2003).

Another possible factor that could explain the low level of variation worldwide is ongoing gene flow over long distances among populations of *C. elegans* worldwide in the form of dauer larva (it is an alternative larval stage in *Caenorhabditis elegans* which allows animals to survive through periods of low food availability during periods of environmental stress) which aids for dispersal by birds, insects, animals and even human (Hodgkin and Doniach 1997).

1.2 Population Structure of *C. elegans*

When the natural isolates are partitioned by world region, most of the genetic variations have been found within a continent for example; the microsatellite genetic diversities within North America and Europe are of 0.83 and 0.85, respectively, which is close to the 0.92 diversity over the whole dataset (Haber et al. 2005). At a local scale,

Haber et al. (2005) and Sivasundar and Hey (2005), found microsatellite polymorphism within compost heaps in Germany, and soil samples in California, respectively. In the French samples too, local populations of *C. elegans* present remarkably high level of polymorphism when compared with diversity found worldwide.

Using AFLP analysis, Barrière and Félix estimated that within-population nucleotide diversity ranged from 0.02 to 0.43 x 10^{-3} per nucleotide for different populations (N_e = 200-5,100). The within-sample diversity is spread over a large range, from almost nil to a level of diversity comparable to the global diversity. In cases where the intra-sample diversity was low, a diversity of 0.80 x 10^{-3} was recovered when three samples were sampled 10-30 meters away (Barrière and Félix, 2005). Thus, local diversity seems to remain high until the scale of a few cm^3. In some samples, it then breaks down, probably as a consequence of local bottlenecks followed by selfing of a single individual. The large variance in within-sample diversity may also reflect a lack of demographic stability in some populations, with transient bottlenecks that reduce genetic diversity. Together with known ecological features of *C. elegans*, this suggests that a *metapopulation* dynamics (many small populations, experiencing frequent extinctions and *de novo* colonisations) could apply to its natural populations.

Using different types of microsatellite data Sivasundar and Hey, 2005 found a surprisingly large degree of 'population structure' at the local level (measured by F_{ST}, which compares the average predicted heterozygosity within populations to the overall predicted heterozygosity). For their California samples, the F_{st}=0.826 suggests little gene flow between populations. The French worms showed a high degree of population structure too (F_{st} = 0.44), even for samples just a few meters apart. From the recent

spatio-temporal survey by Barriere and Felix 2007, *C. elegans* populations appeared to have no structure at a very small scale of a few centimetres, and a very strong structure at scales greater than 10m to as far as 300m; at a larger scale of hundreds of kilometres, the structure appears weaker again. Thus although migration is important in fostering local diversity, it does not appear to be sufficient to prevent population structure from emerging at the spatio-temporal scale (Barrière and Félix, 2005). That is, *C. elegans* genotypes are nearly as diverse at this local scale as globally.

1.3 Outcrossing rates & Linkage disequlibrium in *C. elegans*

Besides migration, the peculiar mode of *C. elegans* reproduction undoubtedly influences the pattern of genetic diversity. It is possible that sex ratio in natural populations is influenced both by environment (for example, X-nondisjunction increases with temperature) as have been found in Drosophila (Jeffreys et al. 2003; Grell 1979) and by genes, as various genes affect specifically X-nondisjunction, and different strains have different sex ratios and male fertilities (Hodgkin and Doniach, 1997). Thus, potential variation may exist for selection to finely tune the sex ratio of *C. elegans*.

The big question however concerns the role that outcrossing plays in the population-genetic dynamics of *C. elegans* in nature, and thus whether males are important or not. *C. elegans* males do not mate very efficiently and are even more attracted to *C. remanei* females than to *C. elegans* hermaphrodites (Chasnov and Chow, 2002). On the other hand, mating efficiency increases with hermaphrodite density, and male sperm allow fertilization of many more oocytes than the limited amount of hermaphrodite sperm, thereby increasing hermaphrodite fecundity. There are several

competing hypotheses to explain why males are maintained: Firstly, males are proposed to be non-adaptive vestiges of gonochoristic (male-female) ancestors (Chasnov and Chow, 2002). Secondly, loci involved in producing males could act as selfish genetic elements, though they may be deleterious to hermaphrodite fitness (Cutter et al. 2003). Thirdly, male outcrossing could be adaptive under a variety of possible conditions, such as avoidance of inbreeding depression caused by selfing (though there is no evidence for heterozygote advantage in crosses between divergent inbred strains), a high deleterious genomic mutation rate (but this rate is not high enough by itself to retain outcrossing), or specific ecological conditions (Cutter et al. 2003). This last explanation appears plausible if unstable environments or pathogens select for recombinant variants, thereby carrying along the male genes that promote effective recombination.

Recent studies shed much light on this issue by estimating the frequency of outcrossing in both lab and wild conditions. In laboratory conditions, spontaneous non-disjunction of the X chromosome occurs at a low rate (about 1/1,000), which is subject to genetic variation (Hodgkin and Doniach, 1997) and is altered by environmental factors such as temperature (Nigon, 1949). When N2 males are placed in culture with hermaphrodites at an initial 50% frequency, their proportion is rapidly reduced to the non-disjunction equilibrium, due to a low mating ability and the apparent absence of an advantage for outcrossed over selfed progeny (Hodgkin, 1983; Chasnov and Chow, 2002; Stewart and Phillips, 2002; Cutter and Payseur, 2003).

Outcrossing in the wild on the other hand has some contrasting estimates ranging from .01 – 20%, a discrepancy risen due to the varying ways of measuring it – typically contrasting values obtained by measuring heterozygote frequencies from that measured

from linkage disequilibrium values. Direct estimates of outcrossing rate within local populations measuring heterozygote frequency vary between 2% (Barriere and Felix 2005) to 20% (Sivasundar and Hey 2005). In contrast, inferred outcrossing rates based on levels of linkage disequilibrium between chromosomes lead to values on the order of .01%. Some of the difference can be accounted for by the strong population structure (F_{st} = 0.8) found in these local populations of worms (Barriere and Felix 2005; Sivasundar and Hey 2005). If we treat F_{st} as inbreeding co-efficient F_{is} which is a mean reduction in heterozygosity due to non-random mating within a sub-population, then outcrossing rate is given by $(1- F_{is})/ (1 + F_{is})$ (Weir 1990). In this case this reduction in outcrossing is actually effective outcrossing. So a F_{st} = .8, would lead to an outcrossing estimate of .11. As F_{is} becomes higher, effective outcrossing rate drastically reduces. The outcrossing rates calculated from heterozygotes frequency is a measure at a particular time-point whereas the estimate based on linkage disequilibrium is a measure over a longer period of time taking into account population structure, linkage with loci under selection and changes in outcrossing over time and thus could result in underestimation.

Linkage disequilibrium (LD) represents a quantitative measure of non-random association between polymorphic loci in a population. Patterns of LD in a genome are shaped by various processes such as population structure, mating system, recombination and gene flow, retaining signatures of population genetic events that happened in distant past. Within local *C. elegans* populations, linkage disequilibrium was found to be high for randomly chosen loci in the samples from Northwest Germany (Haber et al., 2005), France (Barrière and Félix, 2005) and more recently from Scotland (Cutter 2006).

C. elegans harbours high LD between chromosomes, as well as high LD within chromosomes that appear to decay slowly over genetic distances.

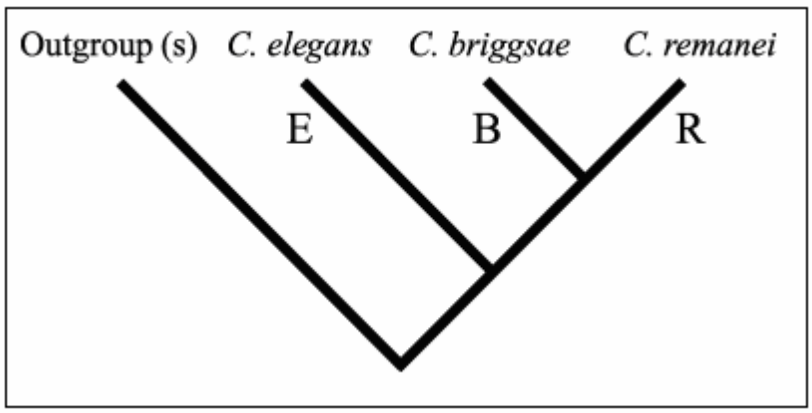
This pattern indicates that little outcrossing maybe occurring in these populations, and consequently very high levels of LD exist, even between unlinked markers. High LD from pooled data may be due to strong population structure among these samples.

1.4 Consequences of a largely selfing population & possible metapulation dynamics

When a population is highly selfing, it can cause the effective outcrossing rate to be much lower than the actual rate at which hermaphrodites fertilise males. As a result, the effective population size (N_e) is reduced and the effect of crossing over on linkage disequilibrium is reduced. This in turn leads to selective sweeps that cause hitchhiking effects over much of the genome. Selfing can also cause an increased effect of background selection (Charlesworth et al., 1993). The cumulative effects of selection and linkage act to further knock down the effective population size. The low outcrossing and low effective population size can lead to accumulation of deleterious mutations thereby increasing the risk of extinction of the population. Such small populations will be inefficient to purge out deleterious mutations and thus can be expected to have a low and fluctuating fitness. The probability of populations going extinct will become high and in turn can lead to turnover of populations with extinction giving way to founding of new populations by migrants from nearby populations that do not yet have a low fitness (Higgins and Lynch 2001; Whitlock and Barton, 1997).Such a population will represent a *metapopulation dynamics* i.e. populations frequently going extinct and habitats being recolonised during migration that will further accelerate genetic drift and reduce

polymorphism levels. Earlier studies on *C. elegans* population genetic structure led to the suggestion that it may follow such a dynamics. Recent evidences from systematic sampling of several locations of natural isolates of *C. elegans* population at different time points clues to this highly dynamic picture. Noticeably, various several putative cases of population extinction events and recolonisation by one or few *C. elegans* individuals of new genotypes have been observed in populations collected from parts of France and Portugal. The temporal scale of population turnover appeared to be of a few weeks or months. And the spatial scale of founding individual migration appears to be ranging from 1 metre to 10 metres. These demographic and genetic results allow us to infer several consequences for genetic and phenotypic evolution of *C. elegans*. The population dynamics of *C. elegans* will thus affect molecular evolution patterns by reducing diversity, increasing LD, and potentially fixing slightly deleterious mutations which maybe compensated at the same or at another locus (Barriere and Felix 2007).

Figure 1.1



Phylogeny for *Caenorhabditis elegans*, *C. briggsae*, and *C. remanei* based on three nuclear loci (globin, *cal-1* and *ceh-13*) – Adapted from Cutter, A. D. & Payseur, B. A. 2003.

Chapter 2

Research design, Materials & Methods:

a. Samples of C. elegans natural isolates:

Soil samples had been collected in the lab (Sivasundar & Hey, 2005) from parks and gardens around Los Angeles area in California (Fig 2.1). From each soil sample that contained *C. elegans*, several lines were established, each starting with a single unmated L4 larval worm. A total of sixty nine lines were established from ten soil samples collected at five locations in California. The soil sample sites were separated by distances ranging from about 10m to about 400m. The lines established from a single soil sample were collectively referred to as population.

In this study, a total of 14 lines were used for gathering Amplified Fragment Length Polymorphisms (AFLP) data which represented 11 populations from 4 different park regions (Table 2.1).

In collections from nature, *C. elegans* co-occurs with other species that are not readily distinguishable from each other based on their morphology, and test-crosses with laboratory strains of *C. elegans* are usually required to determine the species. One of the ways to recover *Caenorhabditis* and some other nematodes from soil or any other substrate like rotten fruits, vegetables, compost heaps is to place the sample onto a standard *C. elegans* culture plate. Alternative methods used by some nematologists have been proven difficult to implement and only a fraction of the individuals could be recovered from such sampling. Species recognition out of the zoo of species that comes with a typical soil or compost sample is also a daunting task. (Barriere and Felix 2006).

Since the occurrence of *C. elegans* is patchy, this also poses difficulty in identifying populations of it in the wild & collecting it. An RNAi collection protocol has thus been developed in the lab (Sivasundar & Hey, 2005) to rapidly identify individual worms. RNAi is a form of post-transcriptional gene silencing where a double stranded RNA silences a gene by specifically binding to its cognate mRNA resulting in a null phenotype for the gene. The gene that was used in the lab was *unc-22* which encodes a muscle protein twitchin which regulates myosin activity. This non-lethal, non-sterile null phenotype is observed relatively quickly characterised by strong twitching. Since the *unc-22* gene in *C. elegans* is only 85% similar to the one in its closest homologue *C. briggsae*, the *C. elegans* sequence is not effective for *C. briggsae* indicating the safe nature of identifying *C. elegans* under this protocol. The accuracy of species identification of the worms isolated by this technique was confirmed through mating tests with lab strains.

b. Cleaning of inbred worm frozen stocks:

As frozen stocks of *C. elegans* were thawed and placed on regular OP50 plates to generate colonies of worms, contaminants of fungi & bacteria were seen growing on the plates. To ensure that the DNA from those contaminants do not interfere with the extraction of worm DNA from the plates, the worm stocks were cleaned following the protocol from Koelle lab (http://info.med.yale.edu/mbb/koelle/protocols/protocol_cleaning_worm_stk.html). Fungal contaminants were cleaned by serial transfer of the worms from one plate to another at regular intervals so that the worms outgrew the mould. Bacterial contaminants were cleaned by subjecting about 10 gravid adults from each line to bleaching treatment

(~.6% NaOCl, 1.5M NaOH). The adults and the bacteria once dead, the newly hatched worms from the eggs were transferred to fresh plates & the lines were thus maintained.

c. DNA extraction:

DNA was extracted from the cleaned worm plates following the protocol of Barriere & Felix, 2005. Worms from recently starved OP50 culture plates were rinsed several times and for several hours in a large volume of M9 buffer to avoid contamination from bacterial DNA if at all there were any residual amount of bacteria present after the cleaning. Genomic DNA was prepared with a Qiagen DNeasy Tissue kit. The concentration of DNA after each isolation was determined using a .8% Agarose-Gel Electrophoresis and also by a DNA fluorometer. Wherever necessary, the final solution was concentrated to get a yield about approximately 100 nanogram per 9 µl of DNA solution. The 100 nanogram of starting concentration forms the starting material for the AFLP analysis.

d. AFLP Protocol:

The AFLP method, a genome wide DNA fingerprinting method, originally outlined and evaluated in detail in the study by Vos et al. (1995) has been used extensively in studying molecular ecology of plants, fungi and bacteria. (Bensch and Akesson 2005). The AFLP method has certain advantages namely – no prior sequence knowledge is needed, repeatability is generally good and the resulting DNA fingerprints provide a large number of genetic markers. The only drawback of this method is that the markers are all dominant and so heterozygotes cannot be distinguished from homozygotes. A brief

overview of the AFLP process has been showed in Fig 2.2 (Mueller and Wolfenbarger 1999). The key steps of the protocol involved the following;

a) Whole genomic DNA was first digested with two restriction endonucleases: *Eco*R I and *Mse* I. *Eco*R I has a 6-bp recognition site GAATTC & *Mse* I has a 4-bp recognition site TTAA. When used together, these enzymes generate small DNA fragments combining all of the three fragments *Eco*R I- *Eco*R I, *Mse* I - *Mse* I & *Eco*R I - *Mse* I. Due to primer design and amplification strategy, the *Eco*R I - *Mse* I fragments are preferentially amplified.

b) These genomic DNA fragments were ligated to *Eco*R I and *Mse* I adaptors that have "sticky ends" to the cut sites opened by the enzymes to generate template DNA for further amplification.

c) The amplification is carried out in two consecutive reactions.

    (i) In the first reaction, called the *preamplification*, the genomic DNAs flanking the regions containing the primer binding site and the restriction site are amplified using two primers, complementary to the adaptor-ligated ends with one preselected nucleotide at the 3'end (in this case *Eco*R I primer containing the selected nucleotide A & *Mse* I primer containing the selected nucleotide C).

    (ii) In the second reaction, called the *selective amplification*, the PCR products of the preamplification are diluted and used as a template using subsequent selective primers having 2 -3 nucleotides at the 3'end of the preamplification primers. By varying the selective bases, it is possible to develop a wide variety of AFLP primer sets, each of which

can reveal a different set of amplified fragments (Vos et al. 1995). The primer combinations used in the following study are given in Table 2.2. The *Eco*R I primer in this case is fluorescent labelled with IRDye 700 for visualising in LI-COR 4200 automated DNA analyser.

(iii) Products from the selective amplification are separated on a 6.5% denaturing polyacrylamide gel electrophoresis. AFLP data as TIF images from IRDye-labelled samples are automatically collected in real time during electrophoresis.

e. Data analyses:

Due to sequence differences in the restriction sites or due to polymorphic indels within the restriction fragments, the polymorphisms are revealed as presence or absence of electrophoretic bands on the gel. A band of certain length represented presence of an allele at one AFLP locus (scored as 1) and populations not having that band of that length indicated absence of the allele (scored as 0) . The gel images obtained from the LI-COR was converted to GelBuddy, a tilling gel analysis tool (Zerr & Henikoff 2005) for proper viewing and detection of the bands. The bands were scored manually & recorded. Though in AFLP, it is impossible to separate dominant homozygous (1/1) from heterozygous (1/0) genotypes, in the case of *C. elegans* for which the strains tend to be highly inbred, with low heterozygosity, the dominance of band presence is not much of an issue. Since the genome sequence for the N2 strain was known, the genomic locations of the bands obtained in AFLP could be directly known from their size. For this, a

program was developed in the lab which could identify the genomic locations of the polymorphisms in the *C. elegans* genome.

The polymorphism data obtained was then analysed for population genetic parameters – Nei diversity (H_j), population structure (F_{ST}) from the AFLP frequencies, following the treatment of Lynch & Milligan (1994). This was done by inputting the data values in AFLPsurv version 1.0 (X. Vekemans, 2002, Universite Libre de Bruxelles, Belgium). The analyses were done assuming an almost completely self-fertilised population, assuming Hardy-Weinberg genotypic proportions.

To calculate pairwise linkage disequilibrium between pairs of alleles, the following metric was used. Linkage disquilibrium (non-random association of alleles) is given by

$D = P_{11}P_{22} - P_{12}P_{21}$, where $P_{11} = p_1q_1$ (gametic frequency of A_1B_1 allele)

$P_{22} = p_2q_2$ (gametic frequency of A_2B_2 allele)

$P_{12} = p_1q_2$ (gametic frequency of A_1B_2 allele)

$P_{21} = p_2q_1$ (gametic frequency of A_2B_1 allele)

This has the disadvantage of depending on the frequency of the alleles inspected. There can be no D observed if any locus has an allele frequency 0 or 1 and is maximal when frequencies are at 0.5. Lewontin (1964) suggested normalising D by dividing it with the theoretical maximum for observed allele frequencies.

Thus $D' = D/D_{max}$, where D_{max} = smaller of p_1q_2 and p_2q_1 if $D < 0$

or D_{max} = smaller of p_1q_1 and p_2q_2 if $D > 0$

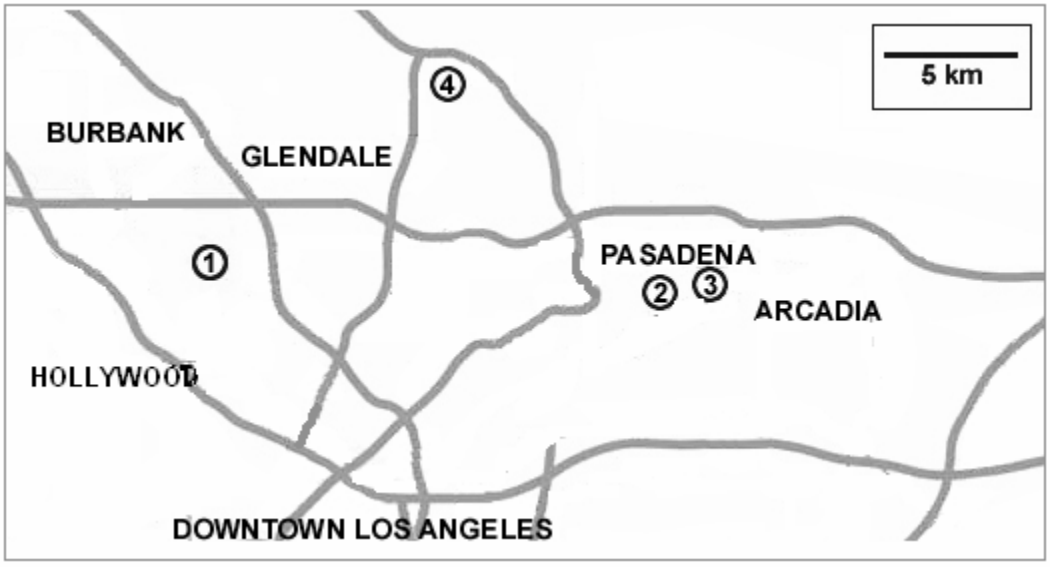
Another measure of linkage disequilibrium used is the correlation coefficient = r^2.

$r^2 = D' / p_1q_1p_2q_2$ (Hill and Robertson 1968)

All the above measures of pairwise linkage disequlibrium were calculated using Microsoft Excel. Tests of the presence of all four combinations of alleles at any two loci (4 gamete test) were performed using Excel with N2 strain included in the data set. The absence of a band was assumed to be homologous, which is a responsible approximation given the low sample wide AFLP diversity.

Figure 2.1

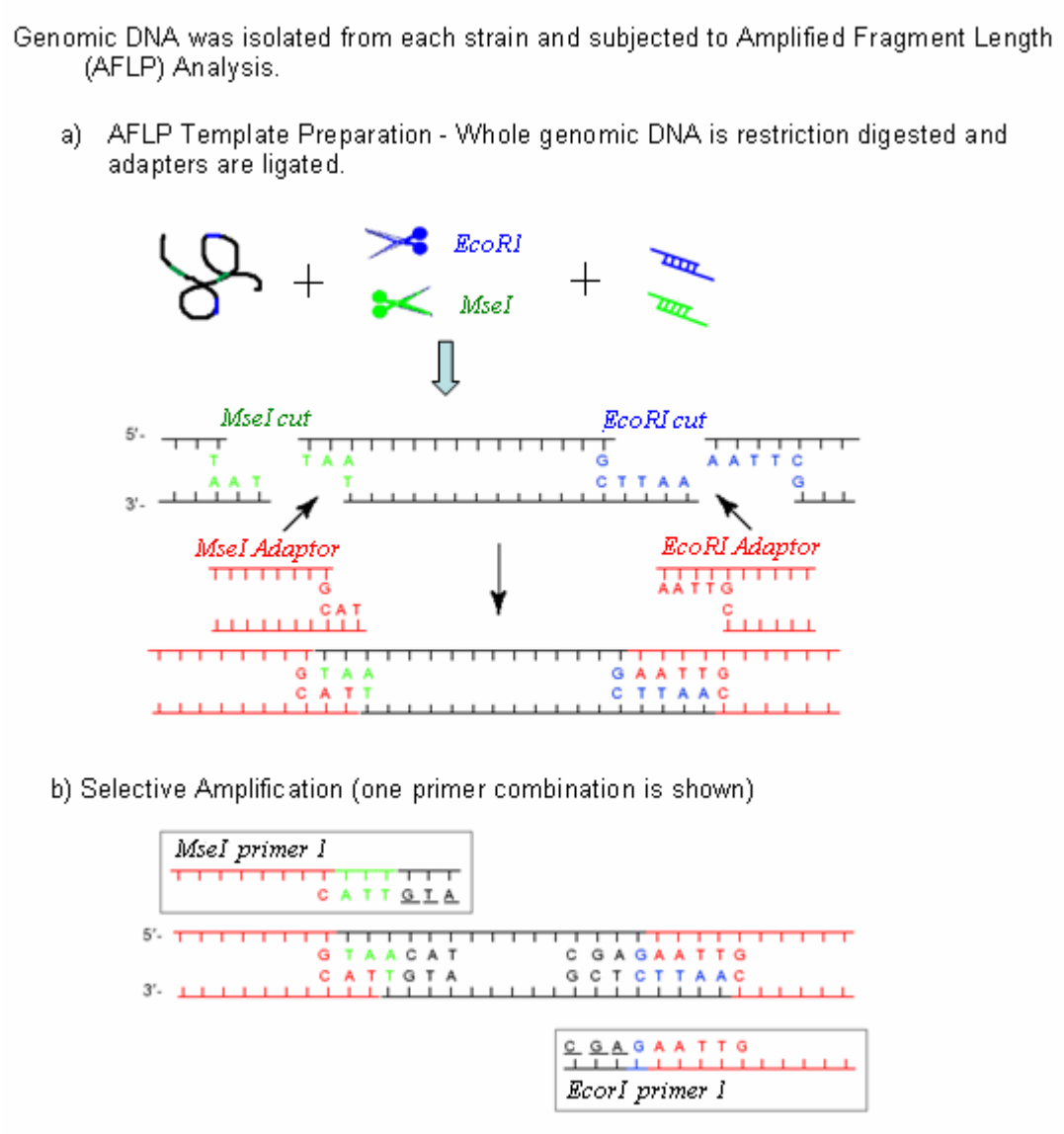
Collection Sites of *C. elegans* wild isolates in the Los Angeles Area of California



Circled numbers indicate the following sample locations: (1) Griffith Park, Los Angeles; (2) Huntington Botanical Gardens, San Marino; (3) California Institute of Technology, Pasadena; (4) Descanso Gardens, La Cañada;

Figure 2.2

AFLP Protocol – Brief Overview



(Figure adapted from Muller & Wolfenbarger, 1999)

Table 2.1

Wild isolates of C. elegans strains used in this study, maintained as frozen stocks in HeyLab

Sl. No	Reference Strain Name	Ref. Population	Location
1.	HEY – 005	w1 Pop	Griffith Park
2.	HEY - 009	w2 Pop	Griffith Park
3.	HEY – 014	GP-4.3 Pop	Griffith Park
4.	HEY – 021	w3 Pop	Huntington Gardens
5.	HEY – 026	w7 Pop	Huntington Gardens
6.	HEY – 027	w7 Pop	Huntington Gardens
7.	HEY – 029	w7 Pop	Huntington Gardens
8.	HEY – 033	w8 Pop	Huntington Gardens
9.	HEY – 040	w11 Pop	Caltech Campus
10.	HEY – 056	w13 Pop	Descanso Gardens
11.	HEY – 082	CT-4.1 Pop	Caltech Campus
12.	HEY – 083	CT-4.2 Pop	Caltech Campus
13.	HEY – 084	CT -4.2 Pop	Caltech Campus
14.	HEY – 086	HG-9 Pop	Huntington Gardens

Table 2.2

List of AFLP Selective primers used in the study

1) *Eco*R I – AT with *Mse* I – CAA
2) *Eco*R I – AT with *Mse* I – CAT
3) *Eco*R I – AT with *Mse* I – CTC
4) *Eco*R I – AT with *Mse* I – CAG
5) *Eco*R I – AT with *Mse* I – CAC

Chapter 3

Results and Discussion:

a. Low genetic diversity indicated by AFLP Polymorphisms:

To analyse genetic variation, sample collection of worms were chosen from worm populations from four locations of California (Griffith Park, Huntington Gardens, Caltech Campus and Descanso Gardens) and from each location, population of worms were collected at distances 10 – 400 metres apart.

Previous studies of DNA sequences and microsatellite variations on these specific strains (Sivasundar and Hey 2005) and also strains collected from worldwide e.g., France, Scotland, Germany found a low overall level of polymorphism in *C. elegans* populations (Barriere and Felix 2005; Cutter 2006; Haber et al. 2005). This fact of low DNA sequence polymorphisms in natural populations of *C. elegans* has led to most of these studies focussing on microsatellite variation. Short Tandem Repeats (STRs) offer the advantages of a greater variability, they can be readily scored, their genomic locations is known and they do not have the ascertainment biases associated with SNP-based studies. However, STRs are not ideal for studies of linkage disequilibrium because of the high mutation rates leading to homoplasy (in which alleles are identical by kind rather than by descent), thereby obscuring pattern of LD. AFLP technique was thus employed, it being less prone to homoplasy than STRs and its ability to detect and score rapidly low levels of polymorphism in a random genome wide manner, without prior knowledge of the sequence.

AFLP was carried on 11 populations of worms collected from 5 location sites in California using a set of 5 different choices of selective primer combinations. As a

reference, the N2 strain of *C. elegans* was also included in the AFLP analysis. Given the small size of the genome, a total of five selective bases were used, fewer than what is used in large plant genomes. A total of 135 fragments were generated in the range of 50 – 300 basepairs in length. One such AFLP profile pattern is shown in Fig 3.1 (after conversion to GelBuddy image from the LI-COR .Tif image) using one of the 5 pairs of primer combinations. Out of the 135 fragments generated over all combinations of primers, 22 polymorphic bands were recorded over the whole data set. This yielded a low overall genetic diversity, with only 22/135 i.e., 16 % of the AFLP fragments being polymorphic over all 11 populations. Almost all of these were shared polymorphisms between one location and the other with only two singleton polymorphism found in one worm population from Caltech Campus and Huntington gardens (Table 3.1). Out of these 22 polymorphisms, 9 were present in the N2 strain and as such their locations on the genome could be predicted based on the genomic sequence of N2.

The overall Nei diversity (H_j) which is defined as the average proportion of loci that is different between pairs of strains (calculated by AFLPSurv 1.0) was estimated to be .0426 i.e., an average of 4.3% of the fragments were divergent between any two genomes sampled at random from the data set. Interestingly, both the estimates of percentage of AFLP polymorphisms & the average Nei diversity for the California strains found in this study was comparable to the ones obtained from a similar study done by Barriere and Felix (2005) from natural isolates of worms collected through regions of France. For the French set of worms, they had found 21% of AFLP fragments to be polymorphic through out their whole data set comprising 149 fragments, yielding a Nei diversity of .049 over their whole data set of worm populations.

b. Substantial Population Structure:

Though selfing has been predominantly the mode of reproduction in the natural populations of *C. elegans* but outcrossing estimates in recent studies have ranged from .01%-20% (Barriere and Felix 2005; Sivasundar and Hey 2005; Cutter 2006). The lower number was estimated from studies on linkage disequlibrium between chromosomes leading to outcrossing estimates of the order of .01%, whereas the higher range was estimated from measures of heterozygote frequencies in populations from France and Los Angeles ranging from 2-20%. But even though rates of outcrossing may be proportional to the rate of male production, the rate of *effective outcrossing* may be much lower. If a male mates with an individual from its same inbred strain, the offspring would be as homozygous as or as heterozygous as would be an individual that arose from self-fertilisation. Effective outcrossing thus occurs when a male has heterozygous offspring by mating with an unrelated individual. Some of the difference in the measurements of outcrossing can be accounted for by the occurrence of strong population structure in the two studies. While using different types of microsatellite data Sivasundar and Hey, 2005 found a surprisingly large degree of population structure at the local level (measured by F_{ST}, which compares the average predicted heterozygosity within populations to the overall predicted heterozygosity) for their California samples, the $F_{st} = 0.826$ suggests little gene flow between populations and strong population structure. The French worms showed a high degree of population structure too at Merlot location ($F_{st} = 0.78$), even for samples just a few meters apart (Barriere and Felix, 2005). Likewise, similar to the observations made from the microsatellite data, the present AFLP study's estimate of

F_{st} = 0.88 for the same population of worms from California indicates a high degree of population structure within local populations of worms. The fact that the polymorphisms found were almost all shared between the locations with only 2 being singleton specific to a single location, reveals the higher local molecular diversity of the worm populations which could have risen due to local mutation or by migration of divergent individuals. But finding of moderately strong F_{st} = 0.44 by pulling all identical multilocus genotypes as a test for spatial structure indicated that similar alleles were found much often within a site than at random (Barriere and Felix, 2005). Taking into account the presence of shared alleles rather than their frequency (reducing the influence of recent selfing on allele frequencies), they found highly significant differentiation between the four locations. Thus although migration is important in fostering local diversity, it might not be sufficient to prevent population structure at the spatio-temporal scale even upto an order of 100 kilometres. The founding of a strong population structure at a smaller scale (10-100m) as well as at a much larger scale (100-1000km) suggests that the *metapopulation* dynamics observed for the worms involve several temporal and spatial scales (Barrier and Felix, 2007) Hence our observation of a large proportion of shared polymorphisms between locations and yet a strong apparent population structure are not incompatible. The sharing of polymorphisms could indicate a continuous influx of migrants or persistence of ancestral polymorphisms for the foundation of the population.

High levels of migration with the aid of humans, flies and birds coupled with strong selection within natural populations can be one possible explanation to reconcile the observations that single populations can possess nearly the same level of variation seen within the species as a whole. But high migrations cannot explain the similarity

among natural isolates because neutral variation would accumulate at introns and intergenic locations. Selection also cannot fully explain the phenomenon, because then one would expect rapid differentiation among local populations at sites not under selection & if the case be so, then selection has to be truly overwhelming to counteract the variation expected to accumulate within local population. An obvious alternative hypothesis to explain the paradoxial nature in worms is then, most of these "natural isolates" have been recently derived from a common source, spread across the globe and then accumulated some local variation. (Phillips 2006).

c. High Linkage disequilibrium measures:

Patterns of linkage disequlibrium which is defined as the non-random association of alleles at different loci (Lewontin and Kojima 1960) are strongly influenced by population genetic processes like inbreeding, gene flow, changes in population size and natural selection. So when a species reveals large numbers of loci in linkage disequibrium, it can be safely taken as a departure from random mating. Linkage disequilibrium both within and between loci have been found to be quite pervasive in *C. elegans* genome. Cutter 2006 found evidence of a strong linkage disequibrium both within and between chromosomes among populations of *C. elegans* from Europe (France, Germany and Scotland) indicating similar ancestries between freely recombining portions of the genome. These results were consistent with patterns observed for microsatellites and AFLPs within the German and French *C .elegans* populations (Barriere and Felix 2005; Haber et al. 2005) and the SNP study of Koch et al. (2000). Unpublished STR data from the Hey Lab including 30 variable loci and 45 wild strains from Caenorhabditis

Genetics Center (CGC) show very high linkage disequlibrium both along the length of Chromosome II as well as among unlinked loci.

LD measures were assessed in the present study using the AFLP data from local Los Angeles population of worms for the 9 polymorphic markers whose genomic location could be identified on the basis of the N2 genome. The relative values for all pairwise comparisons of the different measures of Linkage Disequilibrium – D, D' & r^2 have been shown in Table 3.2. Three of the markers were on X chromosome and were in perfect linkage disequilibrium i.e., D'= 1.0 (Lewontin 1964). The average of the absolute values of LD between these and other markers that were on different chromosomes was much higher than 0, i.e., a mean $|D| = 0.48$.

It is clear that global population of *C. elegans* harbours appreciable LD between chromosomes, as well as high LD within chromosomes that appears to decline slowly over distances. Within a population of constant size, such significant non-random associations between loci can appear by mutation and drift. But, if *C. elegans* temporal dynamics represent a metapopulation mode, then such non-random associations would be expected to be rare. Therefore, this observed linkage disequilibrium must have been present since the foundation of the population, possibly after colonization by the genotypes which would be expected to decay over time as a function of outcrossing rate. In this context, Barriere and Felix' findings in a temporal survey of worms from France of no LD decay over a period of 3 years is significant. If one reconciles this observation as a stochastic effect on populations of small effective population sizes where the sampled populations could be sink populations receiving a constant flow of migrants, each monomorphic for one haplotype, then the absence of decay of LD can be explained

by Wahlund effect (reduction in heterozygosity in a population caused by population substructure – Wahlund 1928). An alternate hypothesis explaining the observation could be selection acting against cross-progeny or recombinants (Barriere and Felix 2007).

Figure 3.1

AFLP Profile

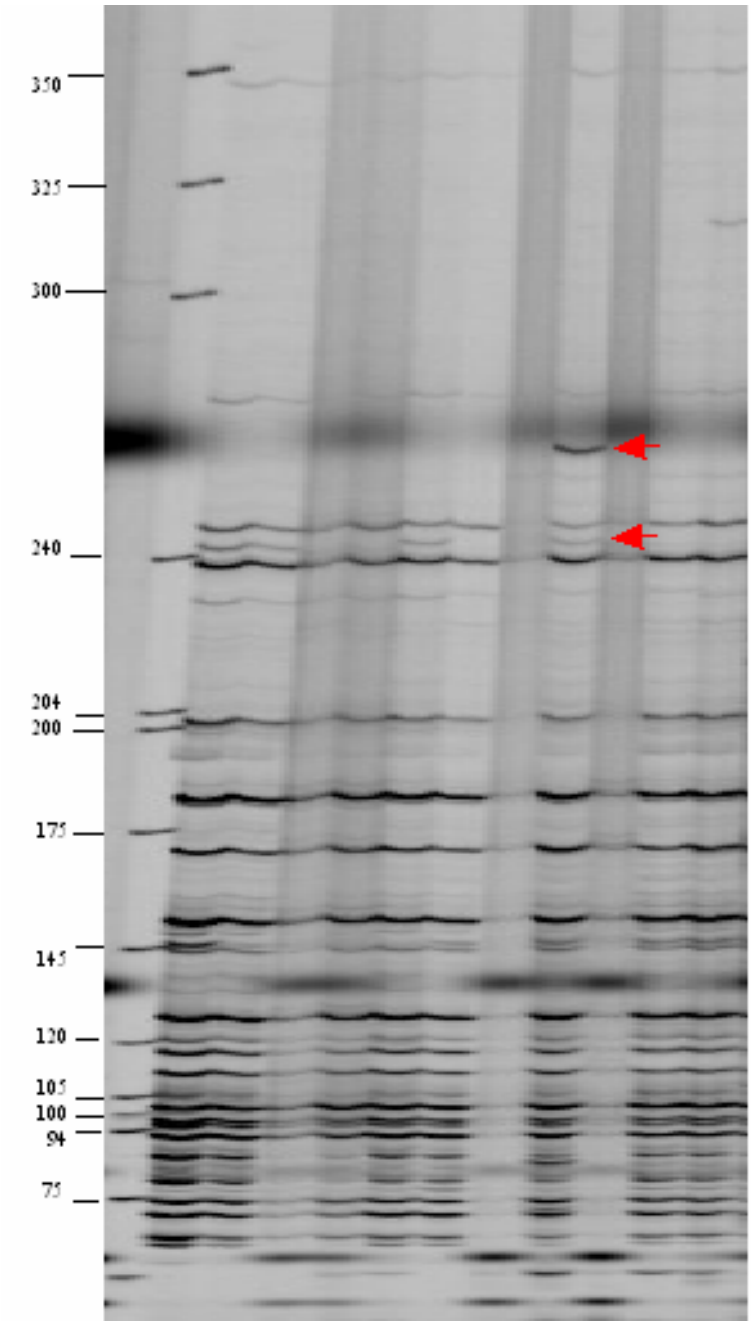


Figure showing a representative AFLP profile using 11 strains of *C. elegans.* The fragments were obtained with primer combination EcoRI -AT & MseI-CAC; characteristic polymorphic bands are marked in red.

Table 3.1

Summary of the AFLP Polymorphisms found

Locations	**Griffith Park**	**Huntington Gardens**	**Caltech Campus**	**Descanso gardens**
Populations Sampled	3	6	4	1
Average Polymorphic Bands	13	14	13	9
Singletons	0	1	1	0

Table 3.2 – Linkage Disequilibrium measures

Table 3.2											
Linkage disequilibrium measures for bands which were found on N2 (to pinpoint their location on the chromosome)											
and are polymorphic within the samples											
			X	X	X	I	V	III	II	IV	V
Measures of D'			**15343887**	**12196887**	**9654246**	**14283636**	**6681615**	**5847228**	**1922331**	**16573838**	**14092619**
			15344031	**12197052**	**9654281**	**14283691**	**6681788**	**5847331**	**1922417**	**16573951**	**14092824**
X	**15343887**	**15344031**		1	1	0.35	0.48	1	0.025	-0.35	0.48
X	**12196887**	**12197052**			1	0.25	0.36	0	-0.25	0	0.125
X	**9654246**	**9654281**				-0.1	1	-1	-1	-1	1
I	**14283636**	**14283691**					0.066667	1	0.066667	-0.06667	-0.06667
V	**6681615**	**6681788**						-1	-0.3	-0.53333	-0.53333
III	**5847228**	**5847331**							1	-1	-1
II	**1922331**	**1922417**								-0.125	-0.19643
IV	**16573838**	**16573951**									0.125
V	**14092619**	**14092824**									
			X	X	X	I	V	III	II	IV	V
Measures of r^2			**15343887**	**12196887**	**9654246**	**14283636**	**6681615**	**5847228**	**1922331**	**16573838**	**14092619**
			15344031	**12197052**	**9654281**	**14283691**	**6681788**	**5847331**	**1922417**	**16573951**	**14092824**
X	**15343887**	**15344031**		0.1875	0.052083	0.034028	0.09	0.052083	0.000625	0.1225	0.09
X	**12196887**	**12197052**			0.454545	0.025	0.098182	0	0.025	0	0.008333
X	**9654246**	**9654281**				0.042735	0.138462	0.005917	0.102564	0.057692	0.102564
I	**14283636**	**14283691**					0.004444	0.138462	0.001852	0.001852	0.001852
V	**6681615**	**6681788**						0.042735	0.066667	0.118519	0.118519
III	**5847228**	**5847331**							0.0625	0.057692	0.0625
II	**1922331**	**1922417**								0.015625	0.038584
IV	**16573838**	**16573951**									0.015625
V	**14092619**	**14092824**									
			X	X	X	I	V	III	II	IV	V
Measures of D			**15343887**	**12196887**	**9654246**	**14283636**	**6681615**	**5847228**	**1922331**	**16573838**	**14092619**
			15344031	**12197052**	**9654281**	**14283691**	**6681788**	**5847331**	**1922417**	**16573951**	**14092824**
X	**15343887**	**15344031**		0.088757	0.029586	0.04142	0.071006	0.029586	0.005917	-0.08284	0.071006
X	**12196887**	**12197052**			0.06944	0.027778	0.070313	0	-0.02778	0	0.020408
X	**9654246**	**9654281**				-0.02551	0.045918	-0.0051	-0.04082	-0.03061	0.04816
I	**14283636**	**14283691**					0.015306	0.045918	0.010204	-0.0102	-0.0102
V	**6681615**	**6681788**						-0.02551	-0.06122	-0.08163	-0.08163
III	**5847228**	**5847331**							0.031111	-0.03061	-0.03111
II	**1922331**	**1922417**								-0.03061	-0.04889
IV	**16573838**	**16573951**									0.030612

Chapter 4

Conclusion:

A species as widespread as *C. elegans* which has a near world-wide occurrence is usually expected to harbour large amounts of genetic variation. However, the finding of low genetic variation along with strong population structure at the local level appears to a paradoxical feature in this partial selfer. Recent studies indicating that local populations of *C. elegans* possessing nearly as much genetic variation as that seen in existing world wide collections is clearly an evolutionary enigma.

Among the various factors that can influence pattern of genetic variation in *C. elegans*, its partial selfing reproductive mode of living is a prominent one. The consequences of a predominantly selfing mode of reproduction is expected to lead to low polymorphism, high linkage disequilibrium and high population subdivision. Also background selection and selective sweeps are expected to be more pronounced under such a mode of reproduction.

As previously noted in earlier studies on worm populations, the findings here also show a low overall genetic diversity in the samples, only 16% of the AFLP fragments being polymorphic. The average Nei diversity H_j = .0426 in the worm samples from California is close to the estimate from similar analyses with *C. elegans* from populations in France (H_j=.049) based on AFLP data (Barriere and Felix, 2005).

Given such estimates of low global diversity, the local genetic diversity had been found to be quite high even at the scale of a few metres. A high degree of population structure is found within local populations; the present study's estimate of F_{st} value of .88 corroborates well with data using microsatellites on these same populations (R_{st} = .86;

Sivasundar and Hey, 2005). The observations are also very similar to the findings within and among the French populations where the Fst values based on AFLP had similar high values (F_{st} = .95 for Merlot population; F_{st} = .45 from pooled data). In addition, high linkage disequlibrium have been found both within and between chromosomes, similar to the recent findings of Cutter (2006). In a similar study using microsatellite data with the same worm populations from California, a high LD was found both along a chromosome (Chromosome II) and also unlinked loci.

The patterns that emerge help us to understand better the evolutionary dynamics of this species over its range. However, they also raise additional intriguing questions presenting investigators with a challenge. Such questions arise like what is the dynamics of genetic variation over time at one location? Are subpopulations of *C. elegans* completely isolated or is there any migration between the subpopulations? Do the linkage disequlibrium within subpopulations and the whole populations indicate population bottlenecks and high rate of local extinction-replacement? Given the likelihood of low outcrossing rates and occurrence of high LD will yield a lower rate of recombination leading to natural selection being compromised. But evidence from patterns of codon usage and gene expression indicates that natural selection is highly effective on individual codons. These may not have easy answers, but the impending work in evolutionary biology using *C. elegans* reaffirms that this worm is undoubtedly a versatile model for addressing fundamental questions of life and can be a useful paradigm.

Bibliography

Abdul Kader, N., and Côté, M. G.1996. Isolement, identification et caractérisation de souches québécoises du nématode *Caenorhabditis elegans*. *Fund. app. Nemat.* **19**, 381–389.

Ayyadevara, S., R. Ayyadevara, S. Hou, J.J. Thaden, and R.J. Shmookler Reis. 2001. Genetic mapping of quantitative trait loci governing longevity of Caenorhabditis elegans in recombinant-inbred progeny of a Bergerac-BO x RC301 interstrain cross. *Genetics* **157:** 655-666.

Ayyadevara, S., J.J. Thaden, and R.J. Shmookler Reis. 2000. Anchor polymerase chain reaction display: a high-throughput method to resolve, score, and isolate dimorphic genetic markers based on interspersed repetitive DNA elements. *Anal Biochem* **284:** 19-28.

Baird, S.E. and W.C. Yen. 2000. Reproductive isolation in Caenorhabditis: terminal phenotypes of hybrid embryos. *Evol Dev* **2:** 9-15.

Barriere, A. and M.A. Felix. 2005. High local genetic diversity and low outcrossing rate in Caenorhabditis elegans natural populations. *Curr Biol* **15:** 1176-1184.

Barrière, A. and Félix,M.-A. Natural variation and population genetics of *Caenorhabditis elegans* (December 26, 2005), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.43.1, http://www.wormbook.org.

Barriere, A. and M.A. Felix. 2007. Temporal Dynamics and Linkage Disequilibrium in Natural Caenorhabditis elegans Populations. *Genetics* **176:** 999-1011.

Bensch, S. and M. Akesson. 2005. Ten years of AFLP in ecology and evolution: why so few animals? *Mol Ecol* **14:** 2899-2914.

Bensch, S., A.J. Helbig, M. Salomon, and I. Seibold. 2002. Amplified fragment length polymorphism analysis identifies hybrids between two subspecies of warblers. *Mol Ecol* **11:** 473-481.

Charlesworth, D. 2006. Population genetics: using recombination to detect sexual reproduction: the contrasting cases of Placozoa and C. elegans. *Heredity* **96:** 341-342.

Charlesworth, B., Charlesworth, D., and Barton, N.H. 2003. The effects of genetic and geographic structure on neutral variation. *Ann. Rev. Ecol. Evol. Syst.* **34**, 99–125.

Chasnov, J.R., and Chow, K.L. 2002. Why are there males in the hermaphroditic species *Caenorhabditis elegans*? *Genetics* **160**, 983–994.

Coghlan, A. and K.H. Wolfe. 2004. Origins of recently gained introns in Caenorhabditis. *Proc Natl Acad Sci U S A* **101:** 11362-11367.

Cutter, A.D. 2004. Sperm-limited fecundity in nematodes: how many sperm are enough? *Evolution Int J Org Evolution* **58:** 651-655.

Cutter, A.D. 2005. Mutation and the experimental evolution of outcrossing in Caenorhabditis elegans. *J Evol Biol* **18:** 27-34.

Cutter, A.D. 2006. Nucleotide polymorphism and linkage disequilibrium in wild populations of the partial selfer Caenorhabditis elegans. *Genetics* **172:** 171-184.

Cutter, A.D., L. Aviles, and S. Ward. 2003a. The proximate determinants of sex ratio in C. elegans populations. *Genet Res* **81:** 91-102.

Cutter, A.D., S.E. Baird, and D. Charlesworth. 2006a. High nucleotide polymorphism and rapid decay of linkage disequilibrium in wild populations of Caenorhabditis remanei. *Genetics* **174:** 901-913.

Cutter, A.D. and B. Charlesworth. 2006. Selection intensity on preferred codons correlates with overall codon usage bias in Caenorhabditis remanei. *Curr Biol* **16:** 2053-2057.

Cutter, A.D., M.A. Felix, A. Barriere, and D. Charlesworth. 2006b. Patterns of nucleotide polymorphism distinguish temperate and tropical wild isolates of Caenorhabditis briggsae. *Genetics* **173:** 2021-2031.

Cutter, A.D. and B.A. Payseur. 2003a. Rates of deleterious mutation and the evolution of sex in Caenorhabditis. *J Evol Biol* **16:** 812-822.

Cutter, A.D. and B.A. Payseur. 2003b. Selection at linked sites in the partial selfer Caenorhabditis elegans. *Mol Biol Evol* **20:** 665-673.

Cutter, A.D., B.A. Payseur, T. Salcedo, A.M. Estes, J.M. Good, E. Wood, T. Hartl, H. Maughan, J. Strempel, B. Wang, A.C. Bryan, and M. Dellos. 2003b. Molecular correlates of genes exhibiting RNAi phenotypes in Caenorhabditis elegans. *Genome Res* **13:** 2651-2657.

Cutter, A.D. and S. Ward. 2005. Sexual and temporal dynamics of molecular evolution in C. elegans development. *Mol Biol Evol* **22:** 178-188.

Cutter, A.D., J.D. Wasmuth, and M.L. Blaxter. 2006c. The evolution of biased codon and amino acid usage in nematode genomes. *Mol Biol Evol* **23:** 2303-2315.

Delattre, M. and M.A. Felix. 2001. Microevolutionary studies in nematodes: a beginning. *Bioessays* **23:** 807-819.

Denver, D.R., S. Feinberg, S. Estes, W.K. Thomas, and M. Lynch. 2005a. Mutation rates, spectra and hotspots in mismatch repair-deficient Caenorhabditis elegans. *Genetics* **170:** 107-113.

Denver, D.R., S. Feinberg, C. Steding, M. Durbin, and M. Lynch. 2006. The relative roles of three DNA repair pathways in preventing Caenorhabditis elegans mutation accumulation. *Genetics* **174:** 57-65.

Denver, D.R., K. Morris, M. Lynch, and W.K. Thomas. 2004. High mutation rate and predominance of insertions in the Caenorhabditis elegans nuclear genome. *Nature* **430:** 679-682.

Denver, D.R., K. Morris, M. Lynch, L.L. Vassilieva, and W.K. Thomas. 2000. High direct estimate of the mutation rate in the mitochondrial genome of Caenorhabditis elegans. *Science* **289:** 2342-2344.

Denver, D.R., K. Morris, J.T. Streelman, S.K. Kim, M. Lynch, and W.K. Thomas. 2005b. The transcriptional consequences of mutation and natural selection in Caenorhabditis elegans. *Nat Genet* **37:** 544-548.

Denver, D.R., K. Morris, and W.K. Thomas. 2003. Phylogenetics in Caenorhabditis elegans: an analysis of divergence and outcrossing. *Mol Biol Evol* **20:** 393-400.

Dolgin, E.S., B. Charlesworth, S.E. Baird, and A.D. Cutter. 2007. Inbreeding and outbreeding depression in Caenorhabditis nematodes. *Evolution Int J Org Evolution* **61:** 1339-1352.

Fitch, D.H. 1997. Evolution of male tail development in rhabditid nematodes related to Caenorhabditis elegans. *Syst Biol* **46:** 145-179.

Fitch, D.H. 2005. Evolution: an ecological context for C. elegans. *Curr Biol* **15:** R655-658.

Fitch, D.H., B. Bugaj-Gaweda, and S.W. Emmons. 1995. 18S ribosomal RNA gene phylogeny for some Rhabditidae related to Caenorhabditis. *Mol Biol Evol* **12:** 346-358.

Fitch, D.H. and W. Sudhaus. 2002. One small step for worms, one giant leap for "Bauplan"? *Evol Dev* **4:** 243-246.

Grell, R.F. 1979. Origin of meiotic nondisjunction in Drosophila females. *Environ Health Perspect* **31:** 33-39.

Haag, E.S., H. Chamberlin, A. Coghlan, D.H. Fitch, A.D. Peters, and H. Schulenburg. 2007. Caenorhabditis evolution: if they all look alike, you aren't looking hard enough. *Trends Genet* **23:** 101-104.

Hartl, D.L., and Clark, A.G. 1997. Principles of Population Genetics (3rd edition) Sunderland, MA, Sinauer Associates, Inc.

Haber, M., M. Schungel, A. Putz, S. Muller, B. Hasert, and H. Schulenburg. 2005. Evolutionary history of Caenorhabditis elegans inferred from microsatellites: evidence for spatial and temporal genetic differentiation and the occurrence of outbreeding. *Mol Biol Evol* **22:** 160-173.

Hedrick, P.W., T.E. Dowling, W.L. Minckley, C.A. Tibbets, B.D. Demarais, and P.C. Marsh. 2000. Establishing a captive broodstock for the endangered bonytail chub (Gila elegans). *J Hered* **91:** 35-39.

Hill, W.G. and A. Robertson. 1968. The effects of inbreeding at loci with heterozygote advantage. *Genetics* **60:** 615-628.

Hillier, L.W., R.D. Miller, S.E. Baird, A. Chinwalla, L.A. Fulton, D.C. Koboldt, and R.H. Waterston. 2007. Comparison of C. elegans and C. briggsae Genome Sequences Reveals Extensive Conservation of Chromosome Organization and Synteny. *PLoS Biol* **5:** e167.

Hodgkin, J. and T.M. Barnes. 1991. More is not better: brood size and population growth in a self-fertilizing nematode. *Proc Biol Sci* **246:** 19-24.

Jeffreys, C.A., P.S. Burrage, and S.E. Bickel. 2003. A model system for increased meiotic nondisjunction in older oocytes. *Curr Biol* **13:** 498-503.

Jovelin, R., B.C. Ajie, and P.C. Phillips. 2003. Molecular evolution and quantitative variation for chemosensory behaviour in the nematode genus Caenorhabditis. *Mol Ecol* **12:** 1325-1337.

Keightley, P.D. and B. Charlesworth. 2005. Genetic instability of C. elegans comes naturally. *Trends Genet* **21:** 67-70.

Kiontke, K., N.P. Gavin, Y. Raynes, C. Roehrig, F. Piano, and D.H. Fitch. 2004. Caenorhabditis phylogeny predicts convergence of hermaphroditism and extensive intron loss. *Proc Natl Acad Sci U S A* **101:** 9003-9008.

Koch, R., H.G. van Luenen, M. van der Horst, K.L. Thijssen, and R.H. Plasterk. 2000. Single nucleotide polymorphisms in wild isolates of Caenorhabditis elegans. *Genome Res* **10:** 1690-1696.

Lamunyon, C.W., O. Bouban, and A.D. Cutter. 2007. Postcopulatory Sexual Selection Reduces Genetic Diversity in Experimental Populations of Caenorhabditis elegans. *J Hered* **98:** 67-72.

Lewontin, R.C. 1964. The Interaction of Selection and Linkage. I. General Considerations; Heterotic Models. *Genetics* **49:** 49-67.

Lynch, M. and V. Katju. 2004. The altered evolutionary trajectories of gene duplicates. *Trends Genet* **20:** 544-549.

Lynch, M. and B.G. Milligan. 1994. Analysis of population genetic structure with RAPD markers. *Mol Ecol* **3:** 91-99.

Maciejowski, J., J.H. Ahn, P.G. Cipriani, D.J. Killian, A.L. Chaudhary, J.I. Lee, R. Voutev, R.C. Johnsen, D.L. Baillie, K.C. Gunsalus, D.H. Fitch, and E.J. Hubbard. 2005. Autosomal genes of autosomal/X-linked duplicated gene pairs and germ-line proliferation in Caenorhabditis elegans. *Genetics* **169:** 1997-2011.

Manoel, D., S. Carvalho, P.C. Phillips, and H. Teotonio. 2007. Selection against males in Caenorhabditis elegans under two mutational treatments. *Proc Biol Sci* **274:** 417-424.

Marais, G. and G. Piganeau. 2002. Hill-Robertson interference is a minor determinant of variations in codon bias across Drosophila melanogaster and Caenorhabditis elegans genomes. *Mol Biol Evol* **19:** 1399-1406.

Mueller, U.G. and L.L. Wolfenbarger. 1999. AFLP genotyping and fingerprinting. *Trends Ecol Evol* **14:** 389-394.

Nei, M., P. Xu, and G. Glazko. 2001. Estimation of divergence times from multiprotein sequences for a few mammalian species and several distantly related organisms. *Proc Natl Acad Sci U S A* **98:** 2497-2502.

Nigon, V. (1949). Modalités de la reproduction et déterminisme du sexe chez quelques nématodes libres. *Ann. des.Sc. Nat. Zool.* ***11***, 1–132.

Phillips, P.C. 2006. One perfect worm. *Trends Genet* **22:** 405-407.

Pires-Dasilva, A. 2007. Evolution of the control of sexual identity in nematodes. *Semin Cell Dev Biol*.

Schulenburg, H. and J.J. Ewbank. 2004. Diversity and specificity in the interaction between Caenorhabditis elegans and the pathogen Serratia marcescens. *BMC Evol Biol* **4:** 49.

Sivasundar, A. and J. Hey. 2003. Population genetics of Caenorhabditis elegans: the paradox of low polymorphism in a widespread species. *Genetics* **163:** 147-157.

Sivasundar, A. and J. Hey. 2005. Sampling from natural populations with RNAI reveals high outcrossing and population structure in Caenorhabditis elegans. *Curr Biol* **15:** 1598-1602.

Teotonio, H., D. Manoel, and P.C. Phillips. 2006. Genetic variation for outcrossing among Caenorhabditis elegans isolates. *Evolution Int J Org Evolution* **60:** 1300-1305.

Vekemans, X., Beauwens, T., Lemaire, M., and Roldan-Ruiz, I. 2002. Data from Amplified Fragment Length Polymorphisms show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Mol eco.***11,** 139-151

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and et al. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* **23:** 4407-4414.

Zerr, T. and S. Henikoff. 2005. Automated band mapping in electrophoretic gel images using background information. *Nucleic Acids Res* **33:** 2806-2812.