The Effects of Breeding Systems on Genetic Architecture

Elie Dolgin

PhD

The University of Edinburgh

2008

Declaration

This thesis was composed by myself, and is a presentation of my original research work, except where otherwise stated in the text.

The work was done under the guidance of Professor Brian Charlesworth at the University of Edinburgh. It has not been submitted for any other degree or professional qualification.

Elie Dolgin, January 31, 2008.

Acknowledgements

This work would not have been possible without the help of a great number of people. Firstly, I would like to thank my supervisor, Brian Charlesworth, for all the stimulating discussions, helpful guidance, and useful comments on many manuscripts and thesis chapters. I am also indebted to my honorary second supervisor, Asher Cutter, who introduced me to the world of the worm, taught me numerous techniques, was lightning-quick with comments on manuscripts, and always had kind words of encouragement.

I am grateful to many members of the Charlesworth lab group for useful discussions, particularly Andrea Betancourt, who was always available to bounce ideas off of or help trouble-shoot problems in the lab. Peter Keightley, Deborah Charlesworth, Laurence Loewe, Kelly Dyer, Penny Haddrill, and Dan Halligan also provided valuable assistance and feedback. Thanks also to the Ashworth support staff, especially Carole Ferrier.

Many members of the worm community generously discussed unpublished results and provided strains, including Scott Baird, Dee Denver, Matt Rockman, Michael Ailion, Karin Kiontke, Henrique Teotónio, and the Caenorhabditis Genetics Center. I would especially like to thank Marie-Anne Félix and Antoine Barrière for teaching me how to identify worms and always providing valuable comments and discussions. I am grateful to Margie Owen-Smith, Alex Müller, Kibidi Amalemba, Charlie Griffiths and Ian Gordon for assistance with field collections in Africa.

Thanks to all my friends in Edinburgh; most of all, Jennifer Carpenter. Finally, thanks to my parents and brother for their long distance support and encouragement.

This work was funded by postgraduate scholarships from the National Science and Engineering Research Council (Canada) and the University of Edinburgh School of Biological Sciences, and tuition fees were paid by an Overseas Research Students Award from Universities UK and an ORS Linked Scholarship from the University of Edinburgh. Various small projects were also funded by the University of Edinburgh Development Trust, the Scottish International Education Trust, and the Genetics Society.

“You have made your way from worm to man, and much in you is still worm.”

– Friedrich Nietzsche in "Thus Spoke Zarathustra"

Abstract

Differences in reproductive strategies are a major factor influencing the patterns of genetic variability. Inbreeding and other non-recombining breeding systems can have profound effects on the efficacy of natural selection, which should be manifested in the patterns of genetic diversity within and between species. The impact of an organism's breeding system can be investigated through a number of approaches. In this thesis, I use mathematical modeling, computer simulations, breeding schemes, quantitative life history measures, and molecular biological techniques to explore many of the consequences of breeding system evolution.

Following a general introduction in Chapter 1, I explore the dynamics of transposable elements (TEs)—selfish mobile sequences of DNA that have deleterious effects upon their hosts. Sexual reproduction and recombination are important for constraining TE abundance, and in the absence of sex, an unchecked proliferation of TEs may cause a population to go extinct. In Chapter 2, I use a theoretical framework to analyze TE dynamics under asexual reproduction. Here, I show that while small populations are driven to extinction by element accumulation, large asexual populations can prevent this fate and be cured of vertically transmitted TEs. These results may help explain an "evolutionary scandal": the persistence of ancient asexual lineages, such as the bdelloid rotifers. In Chapter 3, I extend the computer simulations used in the previous chapter to explore the effects of reduced recombination on the distribution and abundance of TEs in sexual populations. I show that TEs become fixed as a result of Hill-Robertson effects in the form of Muller's ratchet, but only in regions of extremely low recombination when excision is effectively absent and synergism between elements is weak. These results should help explain genomic patterns of TE distributions.

In the remainder of the thesis, I turn to testing the genetic effects of androdioecy—the breeding system in which populations are comprised of separate male and hermaphrodite individuals—using the nematode *Caenorhabditis elegans*

and related species. This unusual breeding system promotes high levels of inbreeding, yet males are maintained at appreciable frequencies. In Chapter 4, I measure life-history traits in the progeny of inbred versus outcrossed *C. elegans* and the related outcrossing species, *C. remanei*, to compare levels of inbreeding depression. I show that highly inbred *C. remanei* show dramatic reductions in brood size and relative fitness compared to outcrossed individuals, whereas pure strains of *C. elegans* performed better than crosses between strains, indicating outbreeding depression. The results are discussed in relation to the evolution of androdioecy and the effect of mating system on the level of inbreeding depression.

Like *C. elegans*, *C. briggsae* reproduces by self-fertile hermaphrodites, and both species have similarly low levels of molecular diversity. But the global sampling of natural populations has been limited and geographically biased. In Chapter 5, I describe the first cultured isolates of *C. elegans* and *C. briggsae* from sub-Saharan Africa, characterize these samples for patterns of nucleotide polymorphism and vulva precursor cell lineage variation, and conduct a series of hybrid crosses in *C. briggsae* to test for genetic incompatibilities. With the new African isolates, I show distinct differences in levels of genetic and phenotypic diversity between the two species. Despite many similarities between *C. elegans* and *C. briggsae*, the results indicate that there may be more subtle, and previously unknown, differences in their natural histories. Finally, I return to the question of the impact of reduced recombination on TE dynamics in Chapter 6, by comparing population frequencies of TEs in natural populations of selfing and outcrossing *Caenorhabditis* species. I show that in the selfing species, *C. elegans*, transposons are less polymorphic and segregate at higher frequencies compared with the outcrossing species, *C. remanei*. Estimates of the intensity of selection based on the population frequencies of polymorphic elements suggest that transposons are selectively neutral in *C. elegans*, but subject to weak purifying selection in *C. remanei*. These results are consistent with a reduced efficacy of natural selection against transposable elements in selfing populations.

Publications

Some of the material in this thesis has already been published in:

- DOLGIN, E. S., and B. CHARLESWORTH, 2006 The fate of transposable elements in asexual populations. *Genetics* **174**: 817-827.
- DOLGIN, E. S., B. CHARLESWORTH, S. E. BAIRD, and A. D. CUTTER, 2007 Inbreeding and outbreeding depression in *Caenorhabditis* nematodes. *Evolution* **61**: 1339-1352.
- DOLGIN, E. S., M.-A. FÉLIX, and A. D. CUTTER, 2008 Hakuna Nematoda: genetic and phenotypic diversity in African isolates of *Caenorhabditis elegans* and *C. briggsae*. *Heredity*, advance online publication 12 December 2007, doi: 10.1038/sj.hdy.6801079.

Contents

1 Introduction 1

1.1 Evolution of Breeding Systems 1

1.1.1 Uniparental Reproduction and Genetic Architecture 2

1.1.2 Inbreeding and Outbreeding Depression 4

1.1.3 Muller's Ratchet 5

1.1.4 Transposable Elements and Restricted Recombination . . . 7

1.1.5 Transposable Elements and Self-Fertilization 8

1.2 *Caenorhabditis* as a Model for Breeding System Evolution 9

1.3 Aims of the Study 12

2 The Fate of Transposable Elements in Asexual Populations 15

2.1 Introduction 15

2.2 Methods 17

2.2.1 Assumptions of the Models 17

2.2.2 Infinite Populations 17

2.2.3 Finite Populations 18

2.2.4 Selection 19

2.2.5 Initial Copy Number of Clonal Lineage 19

2.2.6 Model Parameters 20

2.3 Results 21

2.3.1 Infinite Populations Without Excision 21

2.3.2 Infinite Populations With Excision 24

2.3.3 Parameter Scaling of Finite Populations 25

2.3.4 Finite Populations Without Excision 27

2.3.5 Finite Populations With Excision 29

2.4 Discussion 33

3 The Effects of Recombination Rate on Transposable Elements 39
3.1 Introduction 39
3.2 Methods 42
3.2.1 Population and Genome 42
3.2.2 Simulations 43
3.2.3 Model Parameters 44
2.3 Results 45
2.4 Discussion 51

4 Inbreeding and Outbreeding Depression in *Caenorhabditis* 57
4.1 Introduction 57
4.2 Materials and Methods 60
4.2.1 Nematode Populations 60
4.2.2 Experimental Crosses and Life-History Assays 62
4.2.3 Statistical Analysis 66
4.3 Results 67
4.3.1 Outbreeding Depression in *C. elegans* 67
4.3.2 Inbreeding Depression in *C. remanei* 70
4.4 Discussion 74
4.4.1 Outbreeding Depression in *C. elegans* 74
4.4.2 Implications for *C. elegans* Diversity 76
4.4.3 Frequency of Males 77
4.4.4 Inbreeding Depression in *C. remanei* 77
4.4.5 Evidence for Maternal-Effect Inbreeding Depression 78
4.4.6 Evolution of Androdioecy 79

5 Genetic and Phenotypic Diversity in African Worms 81
5.1 Introduction 81
5.2 Materials and Methods 84
5.2.1 Nematode Populations 84
5.2.2 Molecular Methods 86
5.2.3 Sequence Analysis 87
5.2.4 Vulval Cell Lineage 87
5.2.5 Hybrid Crosses 88

5.3 Results 89
5.3.1 Molecular Polymorphism 89
5.3.2 Vulval Cell Lineage 93
5.3.3 Hybrid Crosses 97
5.4 Discussion 99

6 Transposon Population Frequencies in *Caenorhabditis* 105
6.1 Introduction 105
6.2 Materials and Methods 108
6.2.1 Transposable Elements 108
6.2.2 Nematode Populations and Molecular Methods 109
6.2.3 Population Frequency Assays 109
6.2.4 Estimating Strength of Selection 112
6.3 Results 114
6.3.1 *C. remanei* Transposons 114
6.3.2 Transposon Population Frequencies 114
6.3.3 Strength of Selection 117
6.4 Discussion 119

7 General Discussion and Conclusions 125
7.1 Summary 125
7.2 Future Directions 129

Bibliography 133

1 Introduction

1.1 Evolution of Breeding Systems

A fundamental goal of evolutionary biology is to understand the amount and distribution of genetic diversity within and between populations. One important factor shaping the genetic structure of a population is the breeding system, because how an organism reproduces affects how genes are transmitted from one generation to the next. Evolutionary changes in breeding systems, such as transitions from sexuality to asexuality, or from outcrossing to inbreeding, regularly occur between closely related species. As such, determining the genetic and ecological forces that undermine such changes, as well as the outcomes and consequences of these shifts, can inform on the long-term fate of lineages adopting different reproductive strategies. Although changes in breeding systems have occurred repeatedly in many unrelated genera, selfing or asexual lineages are often evolutionarily short-lived, and the selfing or asexual species we do observe are often of relatively recent origin (MAYNARD SMITH 1992; JUDSON and NORMARK 1996). Why these breeding systems crop up time and again, yet appear ill-suited to survive in the long run, remains an enduring puzzle.

Selfers and asexuals typically have fewer genotypes within populations than randomly mating organisms because there is less genetic admixture among individuals. This "uniparental constraint" means that these populations might be less able to adapt to changing environmental conditions via natural selection (HOLSINGER 2000). Both reproductive modes also reduce a population's effective size (POLLAK 1987), and the size tends to be more variable (SCHOEN and BROWN 1991), making selection less effective, and enhancing the effects of genetic drift. As a result, deleterious mutations that would be eliminated in outcrossing populations will be

more likely to accumulate (KONDRASHOV 1993; WHITLOCK 2002; GLÉMIN 2003). If the build-up of such mutations reduces the population's reproductive potential, the population size can decline further, and make it easier for more harmful alleles to drift to fixation. Such a positive feedback loop could potentially drive the extinction of selfing or asexual populations through a "mutational meltdown" (LYNCH and GABRIEL 1990; GABRIEL *et al.* 1993).

Despite the negative consequences of uniparental reproduction, selfing and asexuality also have their advantages. A major ecological benefit is that these breeding systems bypass the costs associated with finding and courting a mate, which can be especially useful when environmental conditions reduce the opportunity for the union of gametes by two separate individuals, a phenomenon known as reproductive assurance (BAKER 1955). There are also some strong genetic advantages due to the increased transmission of genes to the next generation. Asexual reproduction results in a two-fold advantage because, without any other fitness effects, alleles will spread twice as fast as they would in sexual lineages (MAYNARD SMITH 1978). For selfing, the effect is smaller unless there is an associated increase in allocation to reproductive resources, but there is still predicted to be close to a one-and-a-half-fold advantage (FISHER 1941; CHARLESWORTH 1980). In addition, the faithful nature with which intact genomes are inherited across generations in both breeding systems can constitute an advantage if lineages are well adapted to a particular niche. In this situation, sexual reproduction can break up beneficial gene combinations, thereby imposing a cost known as the recombination load (MASLIN 1968; WHITE 1973).

1.1.1 Uniparental Reproduction and Genetic Architecture

Self-fertilization and asexual reproduction have many similarities in terms of the costs and benefits of restricted genetic admixture, but the direct genetic consequences of the two breeding systems can be markedly different. We can imagine a scenario in which an outcrossing sexual progenitor adopts one of the two uniparental breeding systems. A change to selfing should rapidly increase levels of

homozygosity, thereby exposing recessive or partially recessive alleles to greater selective pressures. Once exposed, advantageous alleles are more likely to spread through the population. This could make inbreeding lineages better able to adapt, compared with outcrossing populations where some degree of dominance is necessary (CHARLESWORTH 1992). But the increased homozygosity is a double-edged sword, and selfing individuals are also more likely to express deleterious mutations and suffer reduced fertility and survivorship. The fitness reduction following mating with close relatives—including one's self—is known as "inbreeding depression" (discussed below).

In contrast to selfers, asexual lineages reproduce their genotypes exactly, except for differences caused by new mutations. So, the frequency of heterozygotes in a diploid or polyploid asexual population can remain quite high, even if there is only a small number of genotypes in the population (BALLOUX *et al.* 2003). In addition, homologous chromosomes can become rearranged, resulting in heteromorphic chromosomes (WHITE 1973). A consequence of long-term heterozygosity in the absence of recombination is that for any locus each allele can acquire different mutations and diverge over time—the so-called "Meselson effect" (JUDSON and NORMARK 1996). Indeed, this criterion has been used to test for ancient asexuality in clonal lineages, such as the bdelloid rotifers (MARK WELCH and MESELSON 2000).

Both uniparental breeding systems also lead to a reduction in the effective recombination rate among loci, resulting in higher rates of linkage disequilibrium. In asexuals or complete selfers, the whole genome is linked and will share the same destiny; even in partially selfing lineages, linkage disequilibrium can still be quite extensive (NORDBORG 2000). Thus, whenever any selected and neutral loci start out in linkage disequilibrium, there should be a decrease in neutral allelic diversity due to the effects of hitchhiking (MAYNARD SMITH and HAIGH 1974) and background selection (CHARLESWORTH *et al.* 1993). Further reductions in diversity are also expected if uniparental reproduction increases the likelihood of population bottlenecks (NEI *et al.* 1975).

1.1.2 Inbreeding and Outbreeding Depression

Over two centuries ago, Thomas Knight noticed that the offspring of self-fertilized garden peas were less vigorous and fertile than outcrossed progeny (KNIGHT 1799). Charles Darwin also pointed out this effect, and devoted an entire book to studying its consequences (DARWIN 1876). Thus, just as the forces promoting the evolution of selfing—reproductive assurance and increased gene transmission—are well established, so is the main factor preventing its spread: inbreeding depression. Inbreeding depression is generally defined as the reduced fitness of inbred individuals compared with their outbred counterparts (WRIGHT 1977). The genetic basis of this fitness decline arises primarily because of increased homozygosity, either for recessive or partially recessive deleterious alleles maintained by recurrent spontaneous mutations, or for alleles at loci where selection favours heterozygotes (CHARLESWORTH and CHARLESWORTH 1999). Evidence for both mechanisms exist in the literature, and the relative importance of each is still under debate, although it is commonly believed that partially recessive mutations play a major role (CHARLESWORTH and CHARLESWORTH 1999; CARR and DUDASH 2003).

Mating systems are expected to have a major impact on the genetic loads that underlie inbreeding depression. If inbreeding depression is due to deleterious recessive alleles, selfing should expose these alleles to selection, and potentially eliminate them from the population (LANDE and SHEMSKE 1985), effectively "purging the genetic load" (CROW 1970). But the degree of purging should be sensitive to the magnitude of the fitness effects of the deleterious alleles. Alleles that are lethal or near-lethal when homozygous should be relatively easily purged, whereas alleles of only small effect will be much more difficult to purge (LANDE and SHEMSKE 1985; CHARLESWORTH *et al.* 1990; HEDRICK 1994; WANG *et al.* 1999). Therefore, different mutations with small deleterious effects can become fixed in different selfing lineages, and crossbreeding should lead to heterosis, or hybrid vigour. This can occur because the effects of deleterious mutations become masked when these mutations find themselves in a heterozygous state. The prediction then follows that selfing populations generally still harbour some inbreeding depression,

although typically at levels lower than that found in outcrossing populations. Indeed, among natural plant populations, inbreeding depression is significantly reduced in self-pollinating species (HUSBAND and SCHEMSKE 1996).

But not all hybrid crosses lead to fitness improvements. When parental genomes are dissimilar beyond a certain point, the progeny of crosses between such parents can result in a fitness decline, known as outbreeding depression (TEMPLETON 1986; LYNCH 1991; WASER 1993; FENSTER and DUDASH 1994). There are several major mechanisms by which conflicting gene interactions can drive this phenomenon. First, if organisms are well adapted to their local conditions, outcrossing among individuals adapted to different environments can produce offspring not adapted to either parental environment. For example, if recessive homozygotes are well suited to a particular environment, as reported for genes responsible for heavy metal tolerance (ANTONOVICS 1968), inbreeding can be favoured. More generally, hybrids grown in parental environments can display poor performance (e.g. SCHMITT and GAMBLE 1990; WASER and PRICE 1994). Second, hybridization between two populations may result in the disruption of co-adapted gene complexes. If the selective advantage of one allele depends on the presence of alleles at other loci, then genetic mixing will break-up beneficial between-locus epistatic interactions (COYNE and ORR 1998; BURKE and ARNOLD 2001). Thirdly, outbreeding depression can result if selection acts against heterozygotes at particular loci (i.e. underdominance; SCHIERUP and CHRISTIANSEN 1996). Lastly, more dramatic chromosomal rearrangements, such as inversions, translocations, or polyploidy, can also affect the compatibility of parental lineages (LEWIS 1973). These mechanisms are not mutually exclusive—for example, genetic coadaptation might underlie adaptation to the local environment (EMLEN 1991; LYNCH 1991).

1.1.3 Muller's Ratchet

Just as the accumulation of mildly deleterious mutations is an expected consequence of recurrent inbreeding (WANG *et al.* 1999), asexual reproduction can also lead to a build-up of deleterious mutations through a process known as

"Muller's ratchet" (MULLER 1964; FELSENSTEIN 1974). In asexual lineages, selection against deleterious mutations will still operate, but if back mutations are rare, when the stochastic loss of all mutation-free individuals in a finite population occurs by genetic drift, they can not be restored and will be lost forever. Thus, the ratchet is deemed to have "clicked" irreversibly. After this click, the class of individuals containing one deleterious mutation becomes the "least loaded" class, which lowers the mean fitness of the population. This class can then be lost in the same way, leading to additional clicks, and the progressive accumulation of deleterious mutations and subsequent fitness degradation can drive the population to extinction. Furthermore, the fixations of deleterious alleles as a consequence of each click can have knock-on harmful effects since the complete linkage between mutations in a non-recombining asexual genomes means that selection is less effective, which is manifested by reductions in the effective population size and lower levels of neutral diversity (GORDO *et al.* 2002).

Escaping the ratchet has been proposed as one of the major reasons for the maintenance of sex (FELSENSTEIN 1974; MAYNARD SMITH 1978; HURST and PECK 1996; BARTON and CHARLESWORTH 1998). Muller's ratchet has also been implicated in the evolution of organelle genomes (LYNCH 1996; LYNCH and BLANCHARD 1998; LOEWE 2006), sex chromosomes (RICE 1994; CHARLESWORTH and CHARLESWORTH 2000), and the degeneration of genomes of obligate symbiont organisms (MORAN 1996; RISPE and MORAN 2000). The rate at which the ratchet clicks is a major factor affecting how important the process should be on an evolutionary scale. Analytic and simulation modeling shows that the ratchet clicks faster when populations are small, mutation rates are high, and selection against the deleterious effects of new mutations are weaker (HAIGH 1978; BUTCHER 1995; GORDO and CHARLESWORTH 2000a,b; STEPHAN and KIM 2002). Therefore, very large asexual lineages with small genomes (i.e. low genomic mutation rates) might persist in the face of Muller's ratchet, but smaller populations with large non-recombining genomes should suffer from its consequences.

1.1.4 Transposable Elements and Restricted Recombination

Although Muller's ratchet has primarily been formulated in terms of classical point mutations, its effects should hold for any major source of mutation, including transposable elements (TEs). TEs are ubiquitous mobile genetic elements found in the genomes of nearly all organisms (CRAIG *et al.* 2002), and there is considerable evidence that TEs cause extensive deleterious effects (CHARLESWORTH *et al.* 1994). Because TEs can self-replicate, they can proliferate even when they incur a substantial fitness cost upon their hosts (HICKEY 1982). This has led to the "selfish DNA hypothesis", which proposes that the abundance of TEs in host genomes reflects a balance between two antagonistic forces: transposition promoting TE copy number, and natural selection acting against their deleterious effects (DOOLITTLE and SAPIENZA 1980; ORGEL and CRICK 1980). But crucially, this balance relies on sexual reproduction. In an asexual lineage, the fate of TEs is entirely linked to the genome of the host. In the long run, TEs that continue to propagate themselves and cause harmful effects can accumulate in a process analogous to Muller's ratchet and potentially drive the extinction of their hosts—and themselves (NUZHDIN and PETROV 2003).

The possibility of TE-mediated lineage extinction has led some authors to argue that selection may favour element-based or host-based attenuation or inactivation of transposition upon the abandonment of sex and recombination, because those lineages with less harmful TEs should be able to outcompete those that have active deleterious TEs (CHARLESWORTH and LANGLEY 1986, 1989; ZEYL and BELL 1996; NUZHDIN and PETROV 2003; ARKHIPOVA and MESELSON 2005a). Consistent with these expectations, cell organelles with uniparental transmission are free of TEs (ZEYL and BELL 1996), and ancient asexual bdelloid rotifers appear to have lost all their deleterious retrotransposons (ARKHIPOVA and MESELSON 2000). These observations have motivated the idea that deleterious TEs could in fact be a primary factor driving the extinction of asexuals, because the rate at which elements could accumulate might be far greater than the rate of Muller's ratchet for classical point mutations (ARKHIPOVA and MESELSON 2005a).

The build-up of TEs in non-recombining genomes is not only expected in asexual lineages. For example, elements also accumulate on the Y-chromosomes of many sexual organisms (GVOZDEV *et al.* 2005; STEINEMANN and STEINEMANN 2005), and recombination does not need to be completely suppressed to promote element accumulation. TEs are often more abundant in chromosomal regions of low recombination (CHARLESWORTH and LANGLEY 1989; DURET *et al.* 2000; BOISSINOT *et al.* 2001; BARTOLOMÉ *et al.* 2002; RIZZON *et al.* 2002, 2003). This pattern is thought to reflect the selective forces limiting TE proliferation in natural populations, and researchers have proposed three major mechanisms by which TEs are constrained: ectopic pairing between non-homologous insertions causing chromosomal rearrangements (MONTGOMERY *et al.* 1991), TEs disrupting gene function (FINNEGAN 1992), and the expression of TE gene products imposing a direct cellular cost (BROOKFIELD 1991). Discriminating between the different models has proven difficult. In the first place, all three mechanisms predict that elements should be more abundant in regions of reduced recombination. Under the first model, if the rate of ectopic exchange is proportional to meiotic exchange, then the number of TEs should be negatively correlated with recombination rate (LANGLEY *et al.* 1988; MONTGOMERY *et al.* 1991; GOLDMAN and LICHTEN 1996, 2000). But the other models both predict that selection against the deleterious effects of insertions should be less efficient in low recombination regions (CHARLESWORTH and LANGLEY 1991). Furthermore, the different mechanisms are not mutually exclusively, and their relative importance remains controversial (BIÉMONT *et al.* 1997; CHARLESWORTH *et al.* 1997; NUZHDIN 1999).

1.1.5 Transposable Elements and Self-Fertilization

Self-fertilization also results in a reduction in the effective recombination rate, but owing to the increased levels of homozygosity, this breeding system can have unique effects on TE dynamics. The outcome of selfing on TEs is particularly sensitive to the mode of selection acting against their deleterious effects (CHARLESWORTH and CHARLESWORTH 1995; WRIGHT and SCHOEN 1999; MORGAN

2001). Under the ectopic exchange model, the deficit of heterozygous insertions in selfing populations may lead to relaxed selection and element proliferation. But if TEs cause direct mutagenic effects, then the TE copy number should decrease under selfing, as selection can more effectively act against insertions in a homozygous state. In addition, other differences between selfing and outcrossing populations can have major impacts on the distribution and abundance of TEs, including changes in effective population sizes and transposition rates.

Comparisons between species with different breeding systems have provided a powerful way to differentiate between the various selective pressures. Although TEs are generally found in low recombination regions of outcrossing species, analyses of the genomes of selfing species have found a relationship between TE abundance and recombination rate for some classes of TEs (DURET *et al.* 2000), but not for most elements (DURET *et al.* 2000; RIZZON *et al.* 2003; WRIGHT *et al.* 2003). Comparisons of TE insertion patterns between closely related self- and cross-pollinating plant species have also produced mixed results. Insertions appear to segregate at higher frequencies in selfing species, but the association between breeding system and differences in abundance is less consistent (WRIGHT *et al.* 2001; TAM *et al.* 2007). Thus, the data indicates a reduction in the efficacy of natural selection against insertions in selfing lineages, but the nature of the relationship between TEs and self-fertilization rates remains unclear.

1.2 *Caenorhabditis* as a Model for Breeding System Evolution

The nematode *Caenorhabditis elegans* is one of the leading model organisms in biological research. Its short generation time, simple structure, ease of laboratory use, ability to reproduce by both selfing and outcrossing, and wealth of genetic and genomic tools established it as a primary tool for molecular biology and genetics. And those same features also make it a compelling organism for studies in evolution, particularly with relation to breeding systems. Long before Sydney Brenner established "the worm project", *C. elegans* and related species were used in comparative studies of reproductive variation and isolation (e.g. MAUPAS 1900;

HONDA 1925; NIGON and DOUGHERTY 1949). However, over the past 40 years most studies with *C. elegans* have been performed on a single common genetic background using the N2 strain, and we know surprisingly little about its natural history. More recently though, wild strains of a number of *Caenorhabditis* species have been isolated, and *C. elegans* and its relatives are emerging as important models for evolutionary biology (DELATTRE and FÉLIX 2001a; HAAG *et al.* 2007).

C. elegans reproduces by a rare and unusual breeding system known as androdioecy, which has only been described for around 50 plants and 36 animals, most of which are nematodes and crustaceans (PANNELL 2002; WEEKS *et al.* 2006a). In *C. elegans*, there are both self-fertilizing protandrous XX hermaphrodites and facultative XO males. Selfing results in hermaphroditic progeny, and occasional males are produced by meiotic non-disjunction of the X chromosome. Outcrossing is only possible through mating with males, resulting in an offspring gender-ratio of approximately 1:1 (HODGKIN *et al.* 1979). This reproductive mode makes *C. elegans* an ideal system to study the evolution of outcrossing versus selfing, the geographic structure of populations and the fate of deleterious mutations. Furthermore, phylogenetic evidence indicates that the ancestor of *C. elegans* was gonochoristic (separate male and female sexes), and that hermaphroditism has evolved at least 10 times, with only one probable case of hermaphroditism reverting to gonochorism (KIONTKE *et al*. 2004; KIONTKE and FITCH 2005). Within the *Caenorhabditis* genus alone, two species independently evolved androdioecy. Therefore, comparisons between related species with the same or contrasting reproductive modes provide a useful platform to understand the genetic changes and ecological forces that promoted the evolution of this atypical breeding system.

The two androdioecious species in the genus, *C. elegans* and *C. briggsae*, were the first pair of closely related animal species to both have their genomes sequenced (*C. ELEGANS* SEQUENCING CONSORTIUM 1998; STEIN *et al.* 2003), and genomic sequencing is currently underway for a number of other species, as well as additional natural isolates. Early gene predictions for *C. remanei* indicate that it contains around 30% more genes than *C. elegans* and *C. briggsae*, and the genome sizes of three gonochoristic species in the genus all appear to be larger than that of the two androdioecous species (HAAG *et al.* 2007). Considering the phylogeny of

Caenorhabditis, these differences suggest that transitions to selfing might consistently result in large reductions in both DNA and gene content, although the selective pressure promoting these reductions remain unclear. The independent origins of selfing also allow inferences about the developmental genetics of parallel evolution. Despite the overtly similar reproductive modes, self-fertility in *C. elegans* and *C. briggsae* has been achieved by using different sets of genes in the core sex determination pathway, indicating a large degree of genetic flexibility in the regulation of breeding systems (HILL *et al.* 2006).

Androdioecy is thought to be unlikely to evolve, because breeding system theory predicts that males must sire at least twice as many offspring as hermaphrodites (LLOYD 1975; CHARLESWORTH 1984). In *C. elegans*, males tend to be driven to low frequencies in laboratory populations (STEWART and PHILLIPS 2002; CUTTER 2005; TEOTÓNIO *et al.* 2006; MANOEL *et al.* 2007). Therefore, there is an ongoing debate about the evolutionary significance of males in *C. elegans* and *C. briggsae*. Because males of most natural isolates are capable of mating (HODGKIN and DONIACH 1997), and rates of protein sequence evolution of male-specific genes are not elevated (CUTTER and WARD 2005), selection seems to maintain male function and outcrossing in the wild. But under laboratory conditions, males and outcrossing events in *C. elegans* are rare (HODGKIN and DONIACH 1997; CHASNOV and CHOW 2002; STEWART and PHILLIPS 2002; CUTTER 2003). Whether this is also true of individuals in the wild is less well known. Studies of molecular variation show *C. elegans* and *C. briggsae* to have very low levels of DNA sequence polymorphism and extensive linkage disequilibrium, especially when compared to the related outcrossing species, *C. remanei* (GRAUSTEIN *et al.* 2002; JOVELIN *et al.* 2003; BARRIÈRE and FÉLIX 2005, 2007; HAAG and ACKERMAN 2005; CUTTER 2006; CUTTER *et al.* 2006a,b). These studies indicate that selfing is the primary mode of reproduction for the androdioecious species, but it has been difficult to estimate outcrossing rates in nature (BARRIÈRE and FÉLIX 2005, 2007; SIVASUNDAR and HEY 2005; CUTTER 2006; CUTTER *et al.* 2006b). The degree with which strains in these species interbreed will directly affect individual homozygosity, effective population sizes, levels of diversity, and population structure (CHARLESWORTH 2003).

Therefore, describing the extent and implications of inbreeding and outbreeding will be crucial to understanding the long-term stability of androdioecy.

1.3 Aims of the Study

This thesis explores the impact of reduced recombination on various aspects of genetic architecture. Recombination and genome mixing are hallmarks of sexual reproduction, and are thought to be primary factors favouring the evolution of sex (MAYNARD SMITH 1978; HURST and PECK 1996; BARTON and CHARLESWORTH 1998; OTTO and LENORMAND 2002). But not all genomes or genomic regions recombine to the same extent. Five different studies have been carried out using a variety of theoretical and empirical approaches to test the effects of differences in effective recombination rates—largely owing to different breeding systems—on levels of genetic variability.

Sexual reproduction and recombination are important for maintaining a stable copy number of TEs (HICKEY 1982). In sexual populations, elements can be constrained by purifying selection against host carriers with higher element copy numbers; however, in the absence of sex and recombination, TEs are bound to their host lineage. Asexual populations could be driven to extinction by an unchecked proliferation of TEs. In Chapter 2, I use mathematical modeling and computer simulation to analyze TE dynamics under asexual reproduction, in order to assess the long term fate of TEs in newly arisen asexual lineages, and explore what might account for the lack of deleterious retrotransposons in the putatively ancient asexual bdelloid rotifers (ARKHIPOVA and MESELSON 2000).

In sexual genomes, TEs often accumulate in regions with suppressed recombination. Whether this pattern reflects a reduction in the efficacy of selection against deleterious insertions or a relaxation of ectopic recombination remains unclear. Discriminating between these two hypotheses has been difficult, because no formal model has investigated the effects of recombination under the model that TEs cause direct deleterious effects. In Chapter 3, I take a simulation-based approach to analyze this scenario and determine the conditions under which element

accumulation is expected in low recombination regions. These results have important implications for differentiating between the leading models of how selection acts on TEs, and should help to interpret emerging population genetic and genomic data.

In Chapters 4–6, I perform three comparative studies using *Caenorhabditis* nematodes to assess the impact of transitions to androdioecy on the genetic structure of populations. Although *C. elegans* reproduces primarily by self-fertilization of hermaphrodites, males are present at low frequencies in natural populations. Males may be maintained in *C. elegans* because outcrossed individuals escape inbreeding depression. The level of inbreeding depression is, however, expected to be low in such a highly selfing species, compared with an outcrosser like *C. remanei.* In Chapter 4, I test for inbreeding depression in life-history traits in the progeny of inbred versus outcrossed *C. elegans* and *C. remanei* individuals derived from recently isolated natural populations. In addition, I maintained inbred lines of *C. remanei* through 13 generations of full-sib mating to assess the effects of prolonged inbreeding.

Investigations of the two selfing species, *C. elegans* and *C. briggsae*, reveal many parallels in terms of morphology, life history, and breeding system. Both species also share similar low levels of molecular diversity. But whereas the distribution of genetic diversity shows a lack of geographic structure in *C. elegans* (DENVER *et al.* 2003; BARRIÈRE and FÉLIX 2005; HABER *et al.* 2005; CUTTER 2006), a clear division is observed in *C. briggsae* between strains derived from different latitudes (GRAUSTEIN *et al.* 2002; CUTTER *et al.* 2006b). However, global sampling of natural populations has been limited mostly to North America and Europe, and it is not clear if the existing samples adequately reflect global diversity. In Chapter 5, I describe patterns of genetic and phenotypic diversity in the first cultured isolates of *C. elegans* and *C. briggsae* from sub-Saharan Africa, and report the results of conducting a series of hybrid crosses in *C. briggsae* to test for genetic incompatibilities. Expanding our understanding of worldwide diversity should help uncover different selective pressures experienced by the two species upon their independent transitions to androdioecy.

Finally, I explore the relationship between self-fertilization and population frequencies of TEs using natural populations of *C. elegans* and *C. remanei*.

Population genetics theory predicts that differences in breeding systems should be an important factor in the dynamics of selfish genetic elements due to different selective pressures experienced both by hosts and by elements. In Chapter 6, I identify a new class of DNA transposon in the *C. remanei* genome, and measure levels of insertion polymorphism of transposons present in the genome sequences of the *C. elegans* N2 strain and *C. remanei* PB4641 strain. This comparison can reveal differences in the efficacy of natural selection against transposable elements in selfing and outcrossing populations.

2 The Fate of Transposable Elements in Asexual Populations

The work presented in this chapter was published in**:**

- DOLGIN, E. S., and B. CHARLESWORTH, 2006 The fate of transposable elements in asexual populations. *Genetics* **174**: 817–827.

Contributing authors:

- I wrote the Mathematica and C++ files, performed the analysis, and wrote the manuscript.
- B. CHARLESWORTH advised on the project, helped with the analytical methods, and assisted in writing the manuscript.

2.1 Introduction

Transposable elements (TEs) are mobile DNA sequences that are abundant in the genomes of nearly all living organisms, including bacteria, protists, fungi, plants, and animals (CRAIG *et al*. 2002). Although there are cases in which mobile genetic elements may be coopted to serve host regulatory or structural functions (KIDWELL and LISCH 2001), like most other mutator mechanisms, TEs are known to reduce the fitness of their host organism. However, unlike other classes of mutation, TEs are capable of autonomous self-replication. When elements transpose, they replicate faster than their host genome, with rates of transposition above rates of spontaneous deletion (CHARLESWORTH and LANGLEY 1989; NUZHDIN and MACKAY 1995; MASIDE *et al*. 2000). This permits TE persistence despite their deleterious effects, as

postulated by the "selfish DNA" hypothesis (DOOLITTLE and SAPIENZA 1980; ORGEL and CRICK 1980).

Experimental and theoretical studies suggest that the ability of TEs to propagate can be balanced by natural selection against individuals with high element copy number (CHARLESWORTH and CHARLESWORTH 1983; KAPLAN and BROOKFIELD 1983). Three main sources of deleterious effects on fitness of segregating TEs have been postulated (see review by NUZHDIN 1999): insertions disrupting gene function (FINNEGAN 1992), chromosomal rearrangements generated by ectopic exchange (MONTGOMERY *et al.* 1987), and a selective cost of transposition itself (BROOKFIELD 1991). Host regulatory mechanisms may play a role as well, and there may be rare beneficial insertions, but it is likely that nearly all TE insertions are deleterious. There is substantial evidence from natural populations and experimental studies for such deleterious effects (CHARLESWORTH *et al.* 1994).

Coevolution between TEs and their hosts depends critically on the presence of sexual reproduction. Outcrossing provides a means for TEs to spread to all individuals in a population. Despite being deleterious, TEs can persist in a sexual population while inflicting a severe fitness penalty on their hosts because they can increase in number faster than the host genome (HICKEY 1982). This would not be true in an asexual population. In the absence of horizontal transmission, there is no between-lineage transmission, and TEs cannot initially spread in an asexual population (HICKEY 1982). However, asexual lineages generally arise from sexual progenitors whose genomes are riddled with TEs. Upon the abandonment of sex, elements are coupled entirely to their hosts. The proliferation of TEs would be detrimental to the host to which they are confined. In the long run, TEs would be expected to become inactive and domesticated, as lineages with inactive elements should have a selective advantage and outcompete other lineages (DOOLITTLE *et al.* 1984). However, theoretical studies show that TEs in outcrossing populations will generally evolve maximum transposition rates (CHARLESWORTH and LANGLEY 1986). Asexual lineages that arise from sexual populations will contain actively transposing TEs, selected to multiply without considering the fitness of the ancestral sexual host. By doing so, they would seal the fates of their asexual hosts. Accordingly, a long-term advantage of sex may result from the early extinction of

asexual lineages, due to the unchecked proliferation of TEs upon the abandonment of sex (ARKHIPOVA and MESELSON 2005a).

Despite the fact that sex facilitates the spread of TEs (HICKEY 1982), it may thus also be necessary to contain their proliferation. In this article, we examine this idea by modeling the population dynamics of transposable elements under asexual reproduction. While the abandonment of sex in small populations will probably lead to element buildup and eventual host extinction in small populations, large populations with some level of element excision or deletion may be able to eliminate vertically transmitted deleterious TEs, providing a potential benefit of asexuality.

2.2 Methods

2.2.1 Assumptions of the Model

We consider a diploid asexual lineage arising from an ancestral sexual population that contains deleterious TEs. With offspring genotypes identical to those of their parents, we assume that a clonal lineage originates from a single ancestor, such that initially all individuals have a TE copy number x. Elements are transmitted only vertically from parent to offspring. The copy number may increase due to transposition, at a rate u per element, or decrease due to excision at a rate v per element. We first consider analytic models of TE dynamics in infinite asexual populations. We then employ a stochastic simulation method to investigate the effects of finite population sizes.

2.2.2 Infinite Populations

To examine TE dynamics with infinite population size, we adapt the exact recurrence relations of KIMURA and MARUYAMA (1966; Equation 3.1), originally devised to calculate the mutation load in asexual populations. We assume an infinite number of insertion sites and Poisson distributions of the numbers of transposition and excision events as an approximation for binomial distributions. We calculate the

transposition load in the presence and absence of excision, without specifying the relation between fitness and copy numbers. We then consider specific fitness functions and determine the mean and variance element of copy number using iterations and approximations based on the method of CHARLESWORTH (1990).

2.2.3 Finite Populations

Computer simulations were used to examine the effects of excision in finite asexual populations. The population size N is constant, with offspring genomes identical to those of their parents except for transposition and excision events. The genome is made up of two diploid chromosomes, with each chromosome able to carry 200 elements, although since the population is asexual and nonrecombining, this is equivalent to any number of chromosomes of any ploidy level. A simulation run is initiated with a single individual carrying x elements, which founds the clonal population, such that all N individuals are initially identical. TE dynamics are then monitored over many generations, where each generation reproduces asexually, followed by transposition and excision. Reproduction involves randomly sampling individuals to produce a new offspring population of size N, where the probability of an individual being selected as a parent is proportional to its fitness. As in the case of the infinite population size model, the numbers of transposition and excision events were drawn from Poisson distributions, with constant probabilities per element of u and v, respectively. New insertions were placed randomly in the genome at unoccupied sites, and the exact locations of elements were maintained between generations unless they had been excised.

Simulations were run until all elements had been lost from the population or until the mean TE copy number had accumulated substantially (>150). Above this number, a runaway process of element buildup is observed (results not shown). No stable equilibrium with an intermediate TE copy number was ever observed, consistent with results from an asexual model of TE sequence evolution for simulations in which TE inactivation through mutation was not included (DOCKING *et al.* 2006). We calculated the proportion of simulation runs where TEs were

eliminated, to examine the conditions under which asexual populations are expected to cure themselves of deleterious TEs or be driven to early extinction. We also monitored the variance in copy number and the minimum TE copy number in the population. C++ files of this simulation program are available upon request. The random numbers were implemented using the Mersenne Twister pseudorandom number generator (MATSUMOTO and NISHIMURA 1998), adapted for C++ by J. BEDAUX (http://www.bedaux.net/mtrand/).

2.2.4 Selection

Fitness was assumed to be a decreasing function of TE abundance, as would be expected in the ancestral population. In sexual populations, a stable equilibrium copy number occurs only when there are synergistic fitness interactions between elements (CHARLESWORTH and CHARLESWORTH 1983). Thus, the fitness of an individual with *n* elements was represented by an exponential quadratic, decreasing function of the copy number,

$$w_n = \exp\left(-an - \tfrac{1}{2}bn^2\right), \tag{2.1}$$

where *a* and *b* are constant selection coefficients (CHARLESWORTH 1990).

2.2.4 Initial Copy Number of Clonal Lineage

We set the initial copy number, *x*, to be equal to the equilibrium mean copy number of the ancestral sexual population. In outcrossing populations, with low frequencies of elements at each occupied site, copy number dynamics are described by

$$\Delta\bar{n} \approx V_n \frac{\partial \ln w_{\bar{n}}}{\partial \bar{n}} + \bar{n}(u - v), \tag{2.2}$$

where $\bar{n}$ is the mean population copy number, V_n is the variance in copy number, and $w_{\bar{n}}$ is the fitness of an individual carrying the mean number of elements (CHARLESWORTH 1985). Solving Equation 2.2 for $\Delta\bar{n} = 0$, assuming a Poisson distribution of elements, and using the quadratic fitness formula, the equilibrium mean element copy number of the ancestral sexual population is

$$\bar{n} \approx (u - a - v)/b. \tag{2.3}$$

The transposition load of a sexual population is measured as

$$L = -\ln \bar{w} = a\bar{n} + \tfrac{1}{2}b\bar{n}^2. \tag{2.4}$$

However, due to the synergism in fitness effects, there is a departure from independence among elements, resulting in the equilibrium load due to TEs being approximately half that of Equation 2.4 when $\bar{n}$ is given by Equation 2.3 (KIMURA and MARUYAMA 1966; CROW 1970; CHARLESWORTH 1990; CHARLESWORTH and BARTON 1996). We thus set the equilibrium load, $L_{eq} = \tfrac{1}{2}L$, and solve for the equilibrium mean copy number under synergistic fitness, such that:

$$n_{eq} = \frac{\sqrt{2\left(a^2 + (u - v)^2\right)} - 2a}{2b}. \tag{2.5}$$

2.2.5 Model Parameters

We set the parameters in our simulations by using estimates from studies of *Drosophila* populations. Average transposition rates per element per generation are ~10^{-4}, regardless of the class of element, with excision rates at least one order of magnitude smaller (NUZHDIN and MACKAY 1995; VIEIRA and BIÉMONT 1997; PASYUKOVA *et al.* 1998; MASIDE *et al.* 2000). We set $u = 10^{-4}$ and investigate excision rates of $v = 10^{-5}$ and $v = 10^{-6}$. The strength of selection on segregating elements in natural populations should be of the order 10^{-5}–10^{-4} per copy, if no

insertions are completely neutral (CHARLESWORTH *et al.* 1994). With synergism between elements, the strength of selection will depend on both a and b. We set $a = 10^{-5}$ and let b vary according to Equation 2.5 for a given n_{eq}. We set the initial copy number, x, of the founding clone to n_{eq}, and assume that asexual lineages will experience the same parameters as the sexual ancestor. We use these values of transposition and selection to investigate the impact of population size on different excision rates.

2.3 Results

2.3.1 Infinite Populations Without Excision

We modify the exact recurrence relations of KIMURA and MARUYAMA (1966; Equation 3.1) to analyze TE dynamics in an infinite asexual population without element excision. We write f_i for the frequency of individuals with i elements, whose fitness is w_i. We assume a Poisson distribution of transposition insertions with mean ui in a genome with i elements, where each transposition event leads to a new insertion. With no excision ($v = 0$), the frequency of individuals with i elements in the next generation is

$$f_i' = \frac{1}{\overline{w}} \sum_{j=0}^{i} f_j w_j \frac{(uj)^{i-j}}{(i-j)!} e^{-uj} , \tag{2.6}$$

where $\overline{w} = \sum f_i w_i$ is the mean fitness. Since there is no excision to reduce the TE copy number below that of the initial clone, the frequency of individuals with the lowest copy number is f_x. Following the derivation of KIMURA and MARUYAMA (1966),

$$f_x' = \frac{f_x w_x}{\overline{w}} e^{-ux}, \tag{2.7}$$

and the lineage reaches an equilibrium with mean fitness

$$\overline{w} = w_x e^{-ux}. \tag{2.8}$$

Since individuals with the lowest element copy number have the highest fitness, all else being equal, a stable equilibrium is achieved with a transposition load,

$$L = ux. \tag{2.9}$$

This is equivalent to the mutational load of an asexual population (KIMURA and MARUYAMA 1966), where the genomic mutation rate, U, is replaced by the product of the transposition rate and initial copy number. Thus, the load depends critically on the initial number of TEs in the founding clone. This result is quite general and independent of the fitness function.

The equilibrium mean and variance in element copy number can be calculated by iteration, given a specific fitness function, transposition rate, and initial copy number. Using Equation 2.1, a normal-distribution approximation can be used to estimate the equilibrium solutions. We adapt Equations 2–4 of CHARLESWORTH (1990), originally developed to calculate the equilibrium number of mutations in an asexual population under mutation–selection balance, to determine the equilibrium properties of transposable elements under asexuality without excision. Using his method of assuming an approximately normal distribution of copy number among individuals in the population, at equilibrium we have

$$\hat{n} = \frac{aV_n}{u + bV_n(u-1)} \tag{2.10}$$

and

$$2\exp\left(-x\left(u + a + \tfrac{1}{2}bx\right)\right) = \ln\left(1 + bV_n\right) - \frac{1}{1 + bV_n}\left(a^2V_n - 2a\hat{n} - b\hat{n}^2\right), \tag{2.11}$$

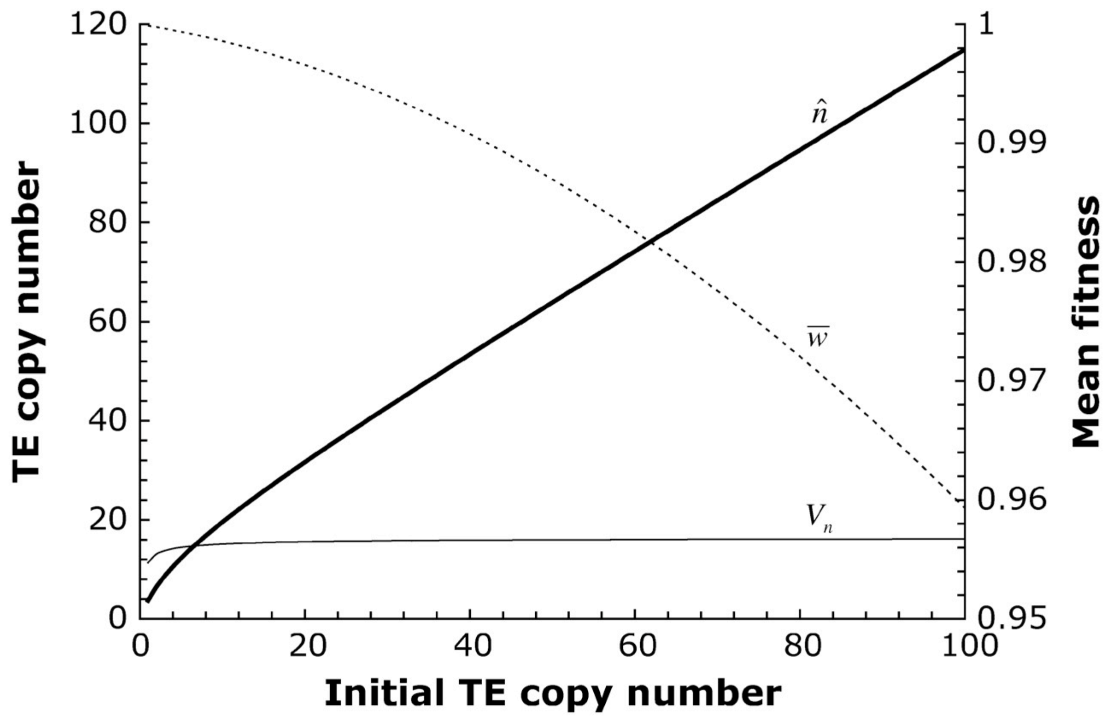


Figure 2.1: Equilibrium mean properties of an asexual population in the absence of excision, at transposition–selection balance with a quadratic fitness function. The primary *y*-axis shows how the mean copy number (thick line) and variance in copy number (thin line) vary with the initial copy number at equilibrium. The dashed line uses the secondary *y*-axis to denote the mean fitness of the equilibrium population. The parameter values used are: rate of transposition per element u = 10^{-4}, selection coefficient a = 10^{-5}, and synergism coefficient b = 6.1 x 10^{-6}.

where $\hat{n}$ is the mean asexual population TE copy number and V_n is the variance in copy number.

These equations allow the mean and variance in TE copy number at equilibrium to be calculated by iteration (CHARLESWORTH 1990). We set $u = 10^{-4}$, $a = 10^{-5}$, $n_{eq} = 10$, and evaluate b using Equation 2.5, giving $b = 6.1 \times 10^{-6}$. We then calculate the equilibrium solutions for an asexual population, with TE copy numbers in the initial clone from 1–100. Figure 2.1 shows the equilibrium properties of an infinite asexual population for the given set of parameters solved, using the normal approximation and confirmed by iteration of Equation 2.6. A stable equilibrium is

always achieved, with a mean copy that depends nearly linearly on x. The mean fitness declines with larger x, while the variance in copy number is nearly constant.

Unlike sexual populations, where the logarithm of fitness must decline faster than linearly with TE copy number (CHARLESWORTH and CHARLESWORTH 1983), a stable equilibrium can be achieved in an asexual population with an arbitrary fitness function. We now consider selection in the absence of synergism (b = 0), as may be the case if these interactions are dependent on the presence of sexual reproduction and recombination (e.g., with meiotic ectopic exchange; LANGLEY *et al.* 1988). Without synergism, the mean and variance in TE copy number simplify to

$$\begin{aligned} \hat{n} &= \frac{2x(a+u)}{a(2-u)} \approx \frac{a+u}{a}x \\ V_n &= \frac{u}{a}\hat{n} \end{aligned} \quad . \qquad (2.12)$$

These equations were confirmed by comparison with the exact equilibrium solutions obtained by iteration of Equation 2.6. It can be seen that the mean and variance scale linearly with the initial TE copy number, x. If the rate of transposition is much greater than the strength of selection against segregating TEs, the mean copy number and variance can be exceedingly high in the absence of synergism. For example, with the parameters used previously ($u = 10^{-4}$; $a = 10^{-5}$), $\hat{n} \approx 11x$ and $V_n \approx 10\hat{n}$. Thus, while a stable equilibrium is theoretically achieved in the absence of synergism between elements, such fitness interactions may be necessary to hold copy number at a biologically relevant equilibrium.

2.3.2 Infinite Populations With Excision

With element excision, the previous analysis no longer holds. The lowest copy number can be reduced below that of the initial clone, and the individuals with fewer TEs should have a selective advantage, leading to a decline in the mean number of elements in the population. Regardless of the TE count of the initial clone, in an infinite asexual population, it is eventually possible for all TEs to be excised, creating a class of individuals without any elements. Unlike the situation with

recurrent mutations, once such a zero class has been generated, it is immune to further transposition, assuming that there is no horizontal transfer. Due to its fitness advantage, this class will spread to fixation, effectively curing the population of deleterious TEs. This was confirmed by modifying Equation 2.6 to include excision and solving by iteration. The mean and variance in TE copy number always decline to zero irrespective of the rates of transposition and excision, the fitness function, or the initial copy number.

2.3.3 Parameter Scaling of Finite Populations

To further investigate whether asexual populations achieve the equilibrium expectations without excision and whether populations with excision can eliminate all their deleterious TEs, computer simulations were used to examine the effects of finite population sizes. Running simulations using the biologically realistic parameters derived from *Drosophila* populations is very time-consuming, because of the large population sizes and small transposition rates, excision rates, and selection coefficients. According to population genetics theory, if evolutionary forces are weak, so that diffusion approximations can be used, the properties of the system are determined by the values of the products of the deterministic forces and the effective population size (EWENS 2004, Chap. 4). If these values are maintained constant, we should expect the same evolutionary outcomes.

We confirmed this expectation by comparing simulations with the parameters from *Drosophila* and an adjusted version of these parameters for $N = 10^4$, such that $Nu = 1$, with the other ratios kept constant as well. To achieve this, we scale up u, v, a, and b one and two orders of magnitude while scaling down N and the numbers of generations by the same amount. The accuracy of this parameter adjustment is demonstrated in Figure 2.2. By reducing the population size and the time of the simulation accordingly, this simplification seems to be reasonably accurate, although perhaps somewhat overestimating the mean copy number, especially with higher excision rates. However, this allows us to perform a greater number of simulation runs in a fraction of the computation time and should be a conservative correction, since it potentially biases our results against finding situations when complete TE

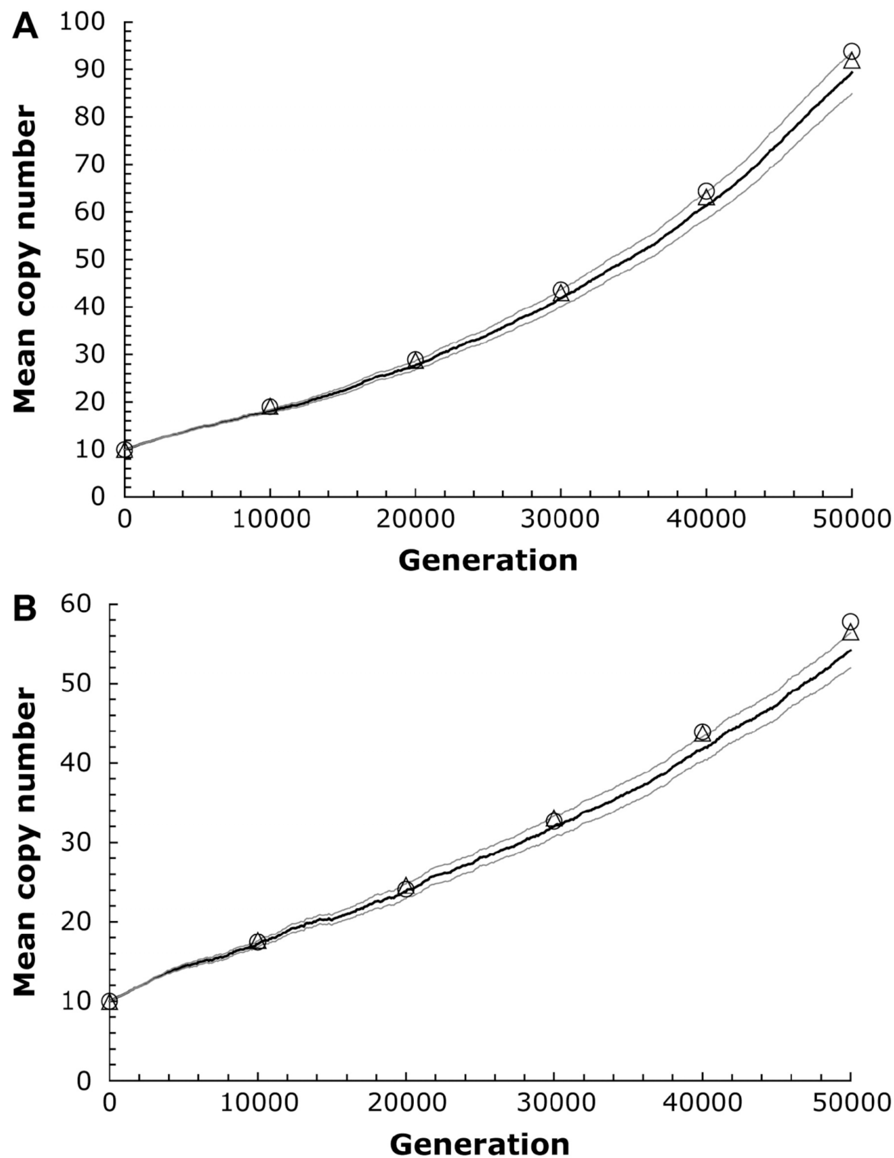


Figure 2.2: Demonstrating the accuracy of scaling parameters. The thick line shows the mean copy number from computer simulations run over 50,000 generations for a population size $N = 10^4$, with initial copy number $x = 10$, rate of transposition per element $u = 10^{-4}$, selection coefficient $a = 10^{-5}$, and synergism coefficient b. The equilibrium is solved for a given rate of excision per element according to Equation 2.5. (A) An excision rate two orders of magnitude lower than transposition ($v = 10^{-6}$; $b = 6.0 \times 10^{-6}$; number of simulation runs, $r = 25$). (B) An excision rate one order of magnitude lower than transposition ($v = 10^{-5}$; $b = 5.4 \times 10^{-6}$; $r = 30$). The thin lines denote the standard error of the mean between runs of the simulation. The triangles show the mean copy number from simulations scaled 10-fold ($r = 100$). The circles show the mean copy number from simulations scaled 100-fold ($r = 1000$). The generation time is also scaled upward to compare with the unadjusted simulation runs.

elimination is expected. We continued by using 100-fold greater transposition, excision, and selection values than those estimated from natural populations. This implies that the population sizes and times used in our simulations need to be scaled upward two orders of magnitude to reflect biologically relevant values. The results presented below show the uncorrected parameter values used in the computer simulations. However, since the magnitude of the population size is so critical to our results, we also include the scaled population sizes in parentheses.

2.3.4 Finite Populations Without Excision

Unlike the case of infinite populations without excision, in simulations with $v = 0$, a large but finite population size, $N = 10^5$ (equivalent to 10^7 when rescaled), and using a scaled set of the parameters presented in Figure 2.1 ($u = 10^{-2}$, $a = 10^{-3}$, and $b = 6.1 \times 10^{-4}$), no equilibrium was ever observed for moderately low initial copy numbers, $x = 5$ and $x = 10$. Elements always accumulated substantially within several thousand generations (up to millions of generations, when rescaled), far above the infinite population equilibrium values of $\hat{n} = 12.3$ and $\hat{n} = 19.5$, respectively. This process of element accumulation is accelerated in smaller populations, with larger initial copy numbers, and with weaker selection. This is consistent with the results for TE sequence evolution in simulations with no excision, where no stable equilibrium in copy number was ever reached (DOCKING *et al.* 2006). This suggests that asexual populations with no excision of TEs are likely to accumulate large numbers of copies within a time frame of millions of generations, even if the effective population size is in the millions. Of course, with a sufficiently large population size, the population will spend a very long time near the equilibrium derived in the previous section, so that it will appear as though no accumulation is happening, but ultimately the population is expected to experience an unbounded proliferation of TEs.

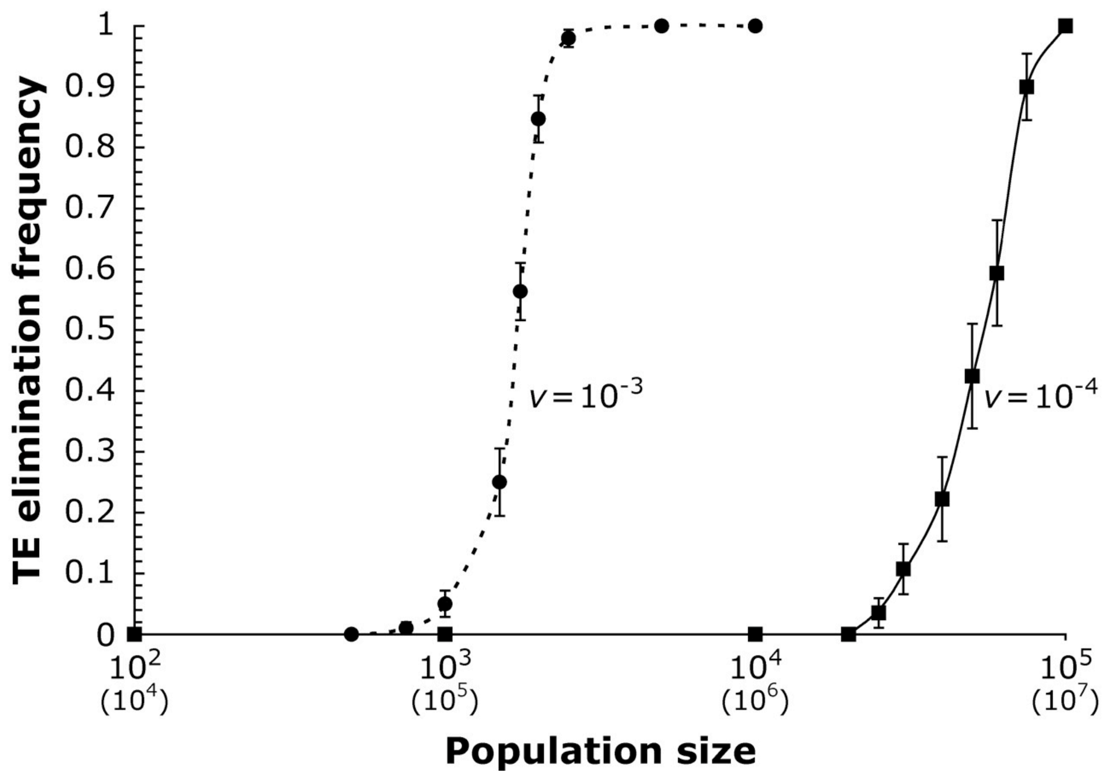


Figure 2.3: TE elimination frequency and population size. The initial copy number was $x = 10$, $u = 10^{-2}$, and $a = 10^{-3}$; the synergism coefficient, b, was calculated for a given rate of excision per element using Equation 2.5. The solid line represents excision two orders of magnitude below transposition ($v = 10^{-4}$; $b = 6.0 \times 10^{-4}$). The dashed line represents excision one order of magnitude below transposition ($v = 10^{-3}$; $b = 5.4 \times 10^{-4}$). With $v = 10^{-4}$, all points with a TE elimination frequency of 0 or 1 are based on at least 10 simulation runs, and all points with an intermediate frequency include at least 30 simulation runs. With $v = 10^{-3}$, all points are based on at least 60 simulation runs. Error bars denote one standard error. To reflect plausible parameters for natural populations, all transposition and selection terms must be scaled down 100-fold. Scaled population sizes are shown in parentheses below the x-axis.

2.3.5 Finite Populations With Excision

As the results of infinite populations with element excision showed, deleterious TEs can be completely purged from asexual populations when $v > 0$. The simulation results illustrate the interaction of population size and excision rate on the probability of TE elimination (Figure 2.3). In contrast to the expectation of complete element elimination in an infinite asexual population, TEs tend to accumulate substantially in small populations. Even with element excision, deleterious TEs accumulate in small populations well above the copy number expected at equilibrium in the absence of excision (see Figure 2.1). On the other hand, in large populations, the efficacy of selection and excision is enhanced. Consistent with the findings for the infinite population model, a sufficiently large asexual lineage can effectively purge itself of vertically transmitting deleterious TEs upon the abandonment of sex.

Two main factors contribute to the proliferation of TEs in small populations. As with sexual populations, this is partly due to the reduction in the variance in copy number between individuals in a finite population, thereby limiting the power of natural selection (BROOKFIELD and BADGE 1997). While we present results only on the final state of the populations, we also monitored the variance in copy number between generations over the course of the simulations. In comparisons among populations that accumulate TEs, smaller populations indeed have a lower variance (results not shown). In addition, random genetic drift may lead to the loss of the class of individuals with the lowest copy number. In the absence of excision, the loss of this class would be irreversible, and TEs would continue to build up, leading to a decline in the mean fitness of the population. This process would be similar to that described for mutations, known as Muller's ratchet (MULLER 1964; STEPHAN and KIM 2002), except that the genomic rate of transposition would increase with the copy number, leading to accelerated rates of fitness deterioration. With element excision, the process is reversible, but the rate at which the least-loaded class is lost may greatly exceed the rate of excision to generate a new best class. We kept track of the minimum TE copy number in the population during our simulations and observed the copy number of the least-loaded class to increase gradually above x as a

population accumulated TEs (results not shown). Both of these factors probably contributed to the increased rate of TE proliferation in smaller populations.

In addition to the rate of excision relative to transposition, the ability of an asexual population to cure itself of deleterious TEs depends critically on the amount of synergism between elements, the magnitude of both transposition and excision, and the initial copy number to be eliminated. Synergism affects the ability of selection to limit TE proliferation in asexual populations. Higher rates of transposition result in faster TE proliferation, even if the ratio of transposition to excision is kept constant. The total number of initial TEs reduces the population mean fitness and affects the total genome-wide rate of transposition. All of these parameters influence the ability of selection and excision to limit copy number.

To explore these factors, we looked at the effect of varying the degree of synergism, b, and changing the magnitude of transposition and excision, while maintaining the ratio of u:v constant at 10. None of these parameters are strictly independent, because they influence the expected equilibrium distribution of elements in the ancestral sexual population. We varied the initial copy number, x, from 10 to 50 and used Equation 2.5 to solve for the corresponding equilibrium value of b and the equilibrium values of u and v. In both cases, for the parameters that were held constant, we used the values employed previously. Figure 2.4 shows the effects of synergism, magnitude of transposition–excision, and initial copy number on the probability of TE elimination at two large population sizes.

All else being equal, increased synergism and lower transposition–excision rates reduce the population size necessary for TE elimination. Increased synergism strengthens selection against large copy numbers, and lower rates of transposition–excision slow the rate of TE accumulation, thereby improving the efficacy of selection. With strong synergism or low rates of transposition–excision, TEs are quite effectively removed from large populations regardless of the initial copy number. In contrast, with low initial copy numbers, TEs are likely to be eliminated, even with weak synergism or strong transposition–excision. Using parameters where intermediate frequencies of TE elimination are observed, a fine balance exists between selection and genetic drift. Slight differences in any of the parameters or the population size can shift this balance and greatly affect whether an asexual

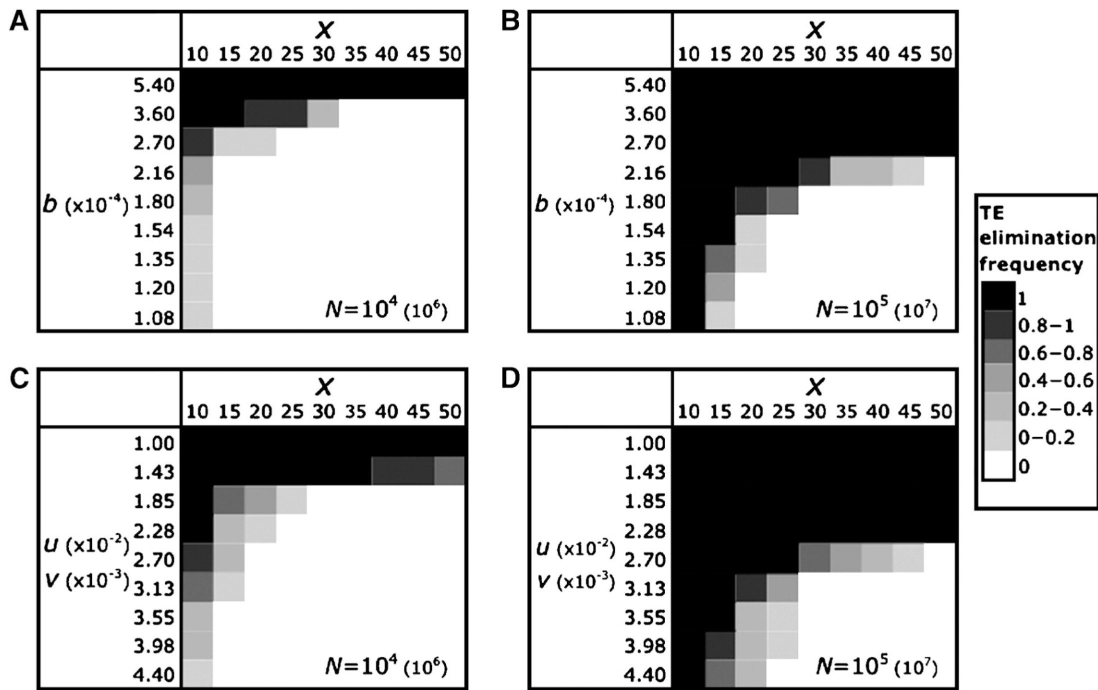


Figure 2.4: The effect of varying the degree of synergism, magnitude of transposition and excision, and initial copy number on the frequency of TE elimination for two large population sizes. (A and B) The effect of changing *b*. (C and D) The effect of changing the magnitude of *u* and *v* while keeping the ratio of *u*:*v* constant at 10, while allowing *x* to vary. The ancestral sexual equilibrium solutions according to Equation 2.5 are located along the diagonal. The parameters that do not change are $u = 10^{-2}$, $v = 10^{-3}$, $a = 10^{-3}$, and $b = 5.4 \times 10^{-4}$. All points are based on at least 10 simulation runs. To reflect plausible parameters for natural populations, all transposition and selection terms must be scaled down 100-fold. Scaled population sizes are shown in parentheses next to the values used in the simulations.

population increasingly accumulates or cures itself of deleterious TEs. Weaker synergism or higher transposition–excision rates also correspond to larger equilibrium mean copy numbers in an ancestral sexual population. Thus, the expected initial copy number will be greater, further hindering the ability of an asexual lineage to purge TEs. However, our approach may be viewed as conservative, since for all analyses other than that in Figure 2.4 we initiated the asexual population with the TE copy number, *x*, set to the sexual equilibrium. Given the variance in copy number in the ancestral population, however, an asexual lineage may arise from lower copy number individuals, thereby reducing *x* and increasing

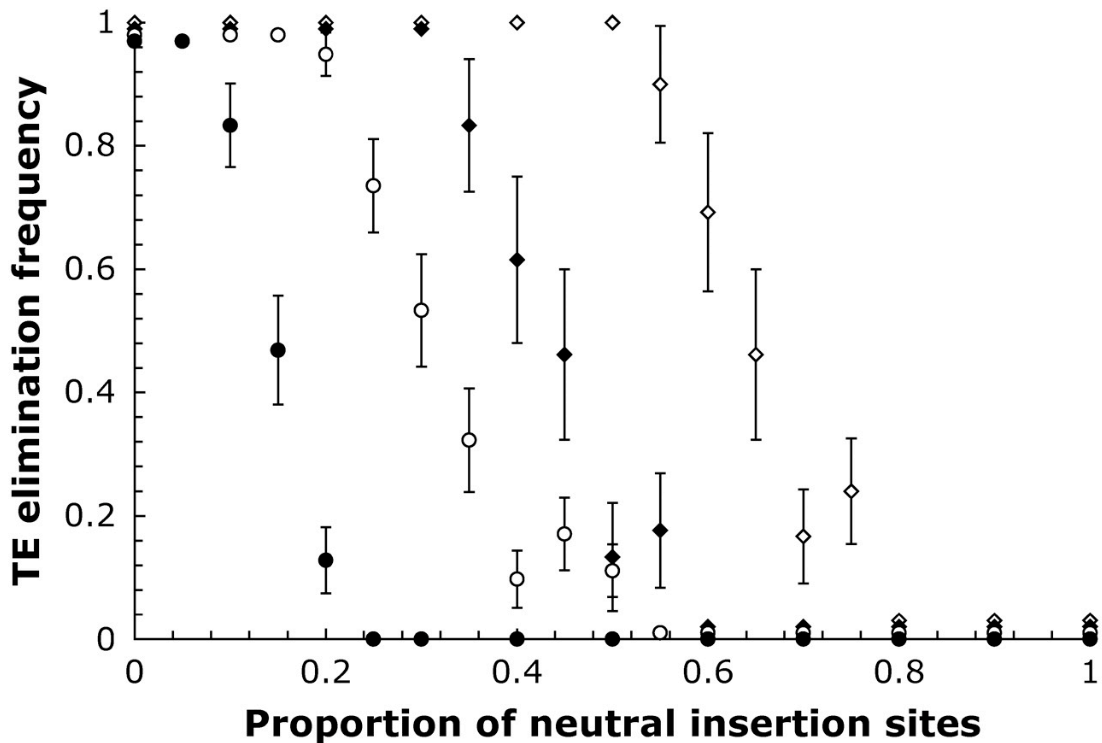


Figure 2.5: TE elimination frequencies incorporating neutral sites. The open symbols show when all x sites containing the initial TEs in the founding clone are under selection; the solid symbols show when these are neutral. The circles show simulations done with $N = 10^4$, diamonds show simulations done with $N = 10^5$, and the remaining parameters are as before: $x = 10$, $u = 10^{-2}$, $v = 10^{-3}$, $a = 10^{-3}$, and $b = 5.4 \times 10^{-4}$. All symbols are based on at least 30 simulation runs for $N = 10^4$ and at least 10 simulation runs for $N = 10^5$. Error bars denote one standard error. To reflect plausible parameters for natural populations, all transposition and selection terms must be scaled down 100-fold, and the population size must be scaled up 100-fold.

the probability of TE elimination. The asexual populations that survive are more likely to have arisen from low copy number founding populations.

Our model has so far supposed that all TE insertions are deleterious, but TEs may vary in their selective effects. In the absence of meiotic ectopic exchange, TE insertions in intergenic regions of the genome may not affect fitness if these regions have little or no functional role. We consider a model with two classes of insertion sites: those in which TEs affect fitness according to Equation 2.1 and those in which TEs are selectively neutral (CHARLESWORTH 1991). We assume that new TE

insertions are neutral with a constant probability p and are deleterious with a probability $1 - p$. Figure 2.5 shows the effect of varying p on the TE elimination frequency. With a high proportion of neutral sites, asexual populations are far less likely to be able to eliminate TEs, since elements tend to build up at neutral sites and continue to transpose. This effect is magnified if TEs inherited from the ancestral sexual population are also present at neutral sites, as would be the case if TEs in this population were primarily contained by ectopic exchange, and hence found in sites where insertions have little or no direct fitness effects in the absence of recombination (MONTGOMERY *et al*. 1987; CHARLESWORTH *et al*. 1994).

2.4 Discussion

Whether an asexual population can cure itself of deleterious TEs depends critically on the ability of selection and excision to limit the propagation of elements. The probability of a population achieving a TE-free state is influenced by the parameters controlling TE dynamics, including the scale of transposition and excision, and the degree of synergism between elements. However, all else being equal, the major factor affecting TE elimination is the population size of the asexual lineage. Since the strength of selection on segregating elements is expected to be relatively low, selection will be ineffective in removing TEs unless the population size is large, and TEs will continue to accumulate at an ever-increasing rate.

Our results are in accordance with the hypothesis that deleterious TEs can drive the extinction of asexuals (ARKHIPOVA and MESELSON 2005a). However, we also provide a means by which an asexual lineage may be able to rid itself of the transposition load that it inherits upon the abandonment of sex. This possibility would provide a long-term benefit to asexual rather than sexual reproduction. Small populations may suffer from a reduced variance in TE copy number and a random loss of the class of individuals with low TE copy numbers. In large populations, however, selection is more effective and the Muller's ratchet-like process of element accumulation can be reversed because of excision. Even rates of TE excision well below that of transposition can improve overall fitness in large populations. Excision

events are doubly advantageous, by both raising fitness and reducing the genome-wide transposition rate because of reduced copy number. This allows mean fitness to improve faster than the stochastic loss of the least-loaded class, leading to the eventual elimination of all TEs. Once a class of individuals free of all TEs has been achieved, they are immune to further transposition and can spread through the population, arresting the process of TE-induced fitness deterioration. But retrotranposons notoriously show little or no excision (CRAIG *et al.* 2002), which poses a problem for this model that we consider below.

Most theoretical models of transposable elements are based on the biology of sexual eukaryotes, as selfish DNA can invade a population only if there is sexual exchange between individuals (HICKEY 1982). The widespread presence of TEs in bacteria has prompted some theoretical analyses of TE dynamics in asexual populations. These models have shown that TE maintenance can be explained by frequent horizontal transmission (SAWYER *et al.* 1987; HARTL and SAWYER 1988), by potential positive impacts of TEs, including beneficial TE insertions causing the fixation of other neutral or slightly deleterious TEs in the genome by hitchhiking (MARTIEL and BLOT 2002), or by selective advantages of TE insertions in fluctuating environments with environment-dependent selection (EDWARDS and BROOKFIELD 2003). BASTEN and MOODY (1991) obtained analytic expressions for equilibrium distributions of TEs in prokaryotic populations, using branching process theory and a number of selection models. In the absence of beneficial insertions, maintenance of TEs requires horizontal transfer, and the net rate of element change due to transposition and excision must be a decreasing function of TE copy number.

Here we focus exclusively on vertically transmitted deleterious TEs in eukaryotic asexual populations. This may be more representative of retroelements, for which horizontal transfer is extremely rare (JORDAN *et al.* 1999; MALIK *et al.* 1999), than of DNA-based transposons, for which there is frequent evidence for horizontal transfer (SILVA *et al.* 2004). Previous authors have proposed that, with negligible transmission rates between individuals, TEs should either accumulate in asexual populations to the point of lineage extinction or be completely eliminated (ARKHIPOVA and MESELSON 2000, 2005a; WRIGHT and FINNEGAN 2001; NUZHDIN and PETROV 2003). However, it remained unclear how asexual lineages could escape

long-term degeneration and completely eliminate all the TEs in the genome. TEs can decay after accumulating mutations, but it seems likely that they will continue to multiply upon the abandonment of sex, given that the rate of inactivating mutations in TEs is probably similar to the rate of transposition (NUZHDIN *et al.* 1998). DOCKING *et al.* (2006) carried out simulations of TE sequence evolution after the abandonment of sex, including the effects of inactivating mutations. They found that, while all elements were eventually lost from asexual populations, it took many thousands of generations for selection on TE sequences to be relaxed, even with higher mutation and transposition rates than expected in nature. This suggests that new asexual lineages should indeed harbour many active TEs, perhaps for millions of generations assuming a scaling of parameters similar to that done in our model.

NUZHDIN and PETROV (2003) suggested that suppressor alleles in a lineage abandoning sex could result in the instantaneous inactivation of TEs. Indeed, our results for infinite populations imply that a lineage with a *trans*-acting repressor (located in either the host or the TE genomes) that reduces the rate of transposition, u, for all elements would have a fitness advantage, by reducing the load due to TEs, $L = ux$ (Equation 2.9). Non-transposing elements would eventually decay under mutation pressure. Our results provide, however, a simpler explanation for complete TE elimination, based on straightforward population genetic processes and biologically plausible parameters and requiring only large population sizes to enhance the efficacy of selection and excision.

The model presented here generally assumes that TEs are uniformly deleterious on a neutral background. The potential for TE elimination may be confounded by the presence of beneficial mutations, whether derived from TEs or otherwise, which may fix linked deleterious TEs by selective sweeps (CHARLESWORTH *et al.* 1992a). However, this requires selection coefficients for beneficial mutations to be substantially greater than those for deleterious TEs (JOHNSON and BARTON 2002; BACHTROG and GORDO 2004). If excision occurs, this would only reduce the rate of TE elimination, by fixing a clonal lineage with a new initial copy number and lower mean fitness. A further limitation on TE elimination arises if a fraction of sites allow neutral TE insertions (CHARLESWORTH 1991). The

continual insertion of TEs into such sites causes a buildup of elements and reduces the ability of selection to contain them at other sites in the genome.

Our model has investigated the eventual fate of an asexual lineage, either complete TE elimination or runaway element accumulation, but has not generally considered the time frame involved. In our simulations, the time to reach either final state generally ranges from several hundred generations to tens of thousands of generations, depending on the parameters involved (results not shown). However, the number of generations generally scales with the population size (see Figure 2.2). This implies that, with the parameter adjustment employed in our model, the time required to drive an asexual lineage to the point of extinction could be as high as millions of generations for larger populations. Similarly, complete TE elimination could take a very long time, especially with high initial copy numbers or weaker selection and excision. Thus, asexual lineages could persist with active deleterious TEs for a substantial period of time before arriving at either final state; however, on a long-term evolutionary scale (i.e., tens of millions of years), it may be the case that only asexual populations that have purged all their deleterious TEs will persist (ARKHIPOVA and MESELSON 2005a).

Asexual lineages generally compose only a single "species" and rarely embrace a taxonomic group of higher rank. While asexual lineages that arise from sexual populations may prosper in the short term, they almost invariably suffer early extinction (see review by JUDSON and NORMARK 1996). Our results have important implications for the existence of ancient asexual taxa: those eukaryotic taxa that have evolved and reproduced for a long stretch of evolutionary time without sex or recombination and have been labeled as "evolutionary scandals" (MAYNARD SMITH 1986). While many claims of ancient asexuality have been refuted upon further investigation, the most compelling evidence for ancient asexuality exists for the bdelloid rotifers, which are not known to have males, hermaphrodites, or meiosis (ARKHIPOVA and MESELSON 2005a). Strikingly, bdelloid rotifers appear to lack deleterious vertically transmitted retrotransposons. Unlike all sexually reproducing eukaryotic species that have been examined, including the monogonont rotifers, reverse transcriptases of two superfamilies of retrotransposons, *gypsy*-like and *LINE*-like retrotransposons, were not detectable in any of five bdelloid species

(ARKHIPOVA and MESELSON 2000). This contrasts sharply with the patchy distribution of various DNA transposons found in bdelloid rotifers (ARKHIPOVA and MESELSON 2000, 2005b), in agreement with their often horizontal mode of transmission, and the presence of a domesticated retrotransposon (ARKHIPOVA *et al.* 2003).

Presumably, the common sexual ancestor of the bdelloid and monogonont rotifers harboured active deleterious TEs, including retrotransposons. It is not certain what mechanism has allowed bdelloids to become free of these TEs. Our models provide a plausible scenario for complete elimination of vertically transmitting TEs, assuming that bdelloids have fairly large population sizes and that excision takes place. It is not known what the effective population size, N_e, of bdelloids might be. This can be estimated from measures of sequence diversity, which is a function of $N_e\mu$, where μ is the mutation rate. While the mutation rate of bdelloid rotifers is unknown, there is no evidence that bdelloids have an unusual mutation rate, as relative rate tests between the nucleotide mutation rates of bdelloid and monogonont rotifers show no significant difference at either synonymous or replacement sites (MARK WELCH and MESELSON 2001). Mitochondrial nucleotide diversity in independently evolving bdelloid clades appears to be similar to that of sexual organisms and averages ~1% (BIRKY *et al.* 2005; C. W. BIRKY, personal communication). Mutation rates of mitochondrial DNA are generally higher than those for nuclear DNA, although this varies between taxa. If we suppose a mutation rate of 10^{-9}–10^{-7}, we can estimate the effective population size to be ~10^5–10^7. This is in the range of being large enough to facilitate complete TE elimination under biologically realistic transposition parameters, even with rates of excision well below that of transposition (see Figures 2.3 and 2.4). Thus, unusually high excision rates or attenuated transposition rates are not necessary to explain the existence of TE-free ancient asexuals.

The transposition, excision, and selection parameters that applied to bdelloid rotifers upon the abandonment of sex are also not known. However, it has been suggested that the unusual ecology of bdelloid rotifers, by causing frequent DNA damage and repair, may have facilitated their loss of TEs (M. MESELSON, personal communication). Bdelloids are capable of anhydrobiosis, a form of dormancy

triggered by desiccation (RICCI 1998). However, unlike most organisms capable of anhydrobiosis, including the monogonont rotifers, bdelloids do not produce the water-replacement substance trehalose (LAPINSKI and TUNNACLIFFE 2003). Bdelloids tolerate high levels of ionizing radiation, causing DNA double-strand breaks (DSBs) that are repaired within a few hours (E. GLADYSHEV and M. MESELSON, unpublished results). The bacterium *Deinococcus radiodurans* is well known for its ability to repair DNA damage following radiation or dehydration by efficient repair mechanisms (BATTISTA 1997), suggesting that the response to irradiation in bdelloids may be an outcome of selection for desiccation resistance.

DSB repair processes in bdelloids might either lead to direct elimination of TEs or cause synergistic selection between deleterious TEs, as a result of ectopic exchange between them (M. MESELSON, personal communication). Although bdelloids do not undergo meiosis, ectopic pairing of TEs during DSB repair could lead to deleterious rearrangements or unrepaired breaks, with a frequency proportional to the square of the copy number (MONTGOMERY *et al.* 1987; CHARLESWORTH *et al.* 1994). Such ectopic pairing may lead to the excision of heterozygous TEs, which is thought to be extremely rare for retrotransposons, and would prevent any TE insertion from being selectively neutral, just as in sexual species, enhancing the possibility of complete TE elimination (see Figure 2.5).

We have shown that newly arisen asexual lineages may be able to avoid early extinction due to TEs. The ability to do so depends critically on large population sizes, with some level of excision. The unique ecology and stress-resistance mechanism of the bdelloid rotifers, which probably evolved as a result of adaptation to their transient aquatic habitats, may have provided the means under which to eliminate deleterious TEs upon the abandonment of sex. Perhaps other asexual lineages were not as fortunate (ARKHIPOVA and MESELSON 2005a).

3 The Effects of Recombination Rate on Transposable Elements

Contributing authors:

- I wrote the C++ files, performed the analysis, and wrote the manuscript.
- B. CHARLESWORTH advised on the project and helped write the manuscript.

3.1 Introduction

Understanding how the proliferation of transposable elements (TEs) is kept in check remains an important theoretical and empirical problem in evolutionary genetics. TEs are widely considered to be intra-genomic parasites, with considerable evidence that they are generally harmful (CHARLESWORTH *et al.* 1994; LYNCH 2007, Chap. 7). But despite their adverse effects, TEs represent a significant proportion of the nuclear DNA of many organisms, and contribute greatly to genome evolution (KIDWELL 2002). The selfish DNA theory posits that the abundance of TEs is regulated by a balance between the tendency for elements to increase in number by transposition and natural selection against the deleterious effects they impose upon their hosts (DOOLITTLE and SAPIENZA 1980; ORGEL and CRICK 1980). However, the nature of the selective forces that limit TE abundance is still poorly understood.

The deleterious effects of TEs have been attributed to three main sources: mutations resulting from insertions into genes or regulatory sequences ("deleterious insertion model"; FINNEGAN 1992); chromosomal rearrangements caused by ectopic recombination between elements in non-homologous insertion sites ("ectopic exchange model"; MONTGOMERY *et al.* 1987); and direct costs due to transposition

activity itself ("deleterious transposition model"; BROOKFIELD 1991). An expected consequence of selection acting to limit element proliferation is that TEs should be more abundant in regions of the genome where they are less likely to be deleterious and/or where natural selection is less effective at removing them. However, part of the problem in discriminating between the different models is that all three make similar and non-mutually exclusive qualitative predictions regarding the genomic distribution of elements. Despite extensive theoretical and experimental results, no general consensus exists on the relative importance of these different factors (BIÉMONT *et al.* 1997; CHARLESWORTH *et al.* 1997; NUZHDIN 1999).

Many studies have tried to test the models by comparing the distribution and abundance of TEs in low versus high recombination regions of the *Drosophila* genome. In the genomes of *Drosophila melanogaster* and *D. buzzatii*, high TE densities have been found in areas of reduced recombination (BARTOLOMÉ *et al.* 2002; RIZZON *et al.* 2002; BERGMAN *et al.* 2006; CASALS *et al.* 2006; FONTANILLAS *et al.* 2007). These genomic data are consistent with *in situ* hybridization data in *Drosophila* that show TEs accumulating in regions where recombination is suppressed, such as the proximal and distal arms of the major autosomes (LANGLEY *et al.* 1988; CHARLESWORTH *et al.* 1992b; MASIDE *et al.* 2001; but see HOOGLAND and BIÉMONT 1996), in polymorphic chromosomal inversions (EANES *et al.* 1992; SNIEGOWSKI and CHARLESWORTH 1994), and on the neo-Y chromosome of *D. miranda* (STEINEMANN and STEINEMANN 1998; BACHTROG 2003). Since areas of low meiotic recombination are expected to have low rates of ectopic recombination (LANGLEY *et al.* 1988; MONTGOMERY *et al.* 1991; GOLDMAN and LICHTEN 1996, 2000), these results may be taken as evidence for the ectopic exchange model.

However, these results are also consistent with the deleterious insertion model, since low recombination regions have lower gene densities (ADAMS *et al.* 2000; FULLERTON *et al.* 2001), and experience weaker effective selection against deleterious insertions (CHARLESWORTH and LANGLEY 1991). Similarly, TE accumulation in low recombination regions is expected under the deleterious transposition model, which makes comparable predictions regarding element patterning to the deleterious insertion model (BROOKFIELD 1991). Therefore, all three models predict a greater abundance of elements in genomic regions of reduced

recombination, and no firm conclusions can be drawn from this observation alone as to whether the accumulation of TEs in low recombination regions reflects differences in the rate of ectopic exchange, or a reduction in the efficacy of selection acting against deleterious element insertions.

While a strong theoretical framework exists for predicting the effects of recombination rate on ectopic exchange and its implications for TE distribution and abundance (LANGLEY *et al.* 1988; CHARLESWORTH *et al.* 1992b), no theoretical studies have taken into account the effects of close linkage on the efficacy of selection under the deleterious insertion model. Local recombination should be an important factor affecting element patterning because a deleterious TE at one site can reduce the efficacy of selection acting at neighboring linked insertion sites, a phenomenon known as the Hill-Robertson effect (HILL and ROBERTSON 1966; FELSENSTEIN 1974). This reduces the effective population size, enhances the importance of genetic drift, and can lead to the fixation of TEs in non-recombining regions through the process of Muller's ratchet (MULLER 1964; FELSENSTEIN 1974). These effects should be influenced by a number of factors, including local recombination rates, effective population sizes, the strength of selection, and the rate of occurrence of new insertions. These factors have been studied for classical mutational models (e.g. GORDO and CHARLESWORTH 2000a, 2001; MCVEAN and CHARLESWORTH 2000; COMERON and KREITMAN 2002; KEIGHTLEY and OTTO 2006), but the unique ability of transposable elements to replicate autonomously through transposition, and excise from the genome, could lead to distinct patterns and evolutionary dynamics. A better understanding of the genomic distribution expected from these effects should help differentiate the deleterious insertion model from the ectopic exchange model.

Theoretical considerations of TEs under the deleterious insertion model have almost universally assumed a constant recombination rate across the entire genome. Part of the reason that the effects of variable recombination have not been well studied is that understanding these processes often requires a simulation-based approach. This is in contrast to the relatively simple prediction of the ectopic exchange model that the number of TE copies inserted in a given region is expected to be inversely related to the recombination rate (LANGLEY *et al.* 1988;

CHARLESWORTH *et al.* 1992b). In this paper, we develop a simulation model that simultaneously considers a genome with both high and low recombination regions. We then identify conditions and parameters under which element accumulation in regions of reduced recombination is expected. We show that the deleterious insertion model can explain an increased abundance of TEs in low recombination areas only when individual element insertions become fixed at the population level. We discuss the results of our simulations in relation to the genomic data and population surveys of *D. melanogaster* to try and distinguish between the leading hypotheses for why TEs build-up in low recombination regions.

3.2 Methods

Computer simulations were used to examine the effects of reduced recombination on the distribution and abundance of transposable elements in the genome. The model was implemented by modifying the simulation program of DOLGIN and CHARLESWORTH (2006) to incorporate sexual reproduction and recombination. C++ files are available upon request. Random numbers were generated using the Mersenne Twister pseudorandom number generator (MATSUMOTO and NISHIMURA 1998), adapted for C++ by J. BEDAUX (http://www.bedaux.net/mtrand/).

3.2.1 Population and Genome

We consider a diploid population of a constant size, *N*, with discrete non-overlapping generations. The genome is composed of four chromosomes: three large recombining chromosomes, each representing 30% of the genome, and one smaller chromosome, representing 10% of the genome. This is intended to provide a rough portrayal of the *D. melanogaster* genome, for which most data on TE distributions are available, although we have exaggerated the size of the small, non-crossing over fourth chromosome (ASHBURNER *et al.* 2004). The larger chromosomes are able to carry as many as 1000 elements each, permitting a large number of potential

insertions, with a recombination frequency, r, of 10^{-3} between any two adjacent potential insertion sites, resulting in one crossover event per chromosome per generation on average. The smaller chromosome can carry a maximum of 333 elements, and the recombination frequency is adjusted in different simulations to analyze the effects of a genomic region with reduced and null recombination, respectively.

3.2.2 Simulations

The simulation is initiated by randomly inserting N elements into the population (i.e. 1 copy per genome on average). Each generation then starts with reproduction, involving selection and recombination, followed by transposition and excision. In each generation, individuals are sampled as follows to produce a new offspring population of size N. For each offspring, two parents are randomly selected. Each parent then creates a gamete, with the number of crossovers per chromosome drawn from a Poisson distribution according to the length and recombination rate of the chromosome, and a uniform distribution of crossover positions. The gametes are then combined, and the resulting offspring are subject to selection, with the probability of survival proportional to their fitness values, as described by DOLGIN and CHARLESWORTH (2006). Fitness is assumed to be a decreasing function of TE abundance. This is essential to allow for a stable equilibrium copy number under free recombination; otherwise, TEs will proliferate in an unbounded fashion regardless of recombination rate (CHARLESWORTH and CHARLESWORTH 1983). We model the fitness of an individual with n elements by an exponential quadratic, decreasing function of the TE copy number,

$$w_n = \exp\left(-an - \tfrac{1}{2}bn^2\right) \tag{3.1}$$

where a and b are constant selection coefficients (CHARLESWORTH 1990), and dominance is assumed to be intermediate at each insertion site (i.e. n is the sum of all the elements in the diploid genome). The model permits TEs to segregate in the population, and implicitly assumes that insertions causing large fitness effects are

rapidly eliminated by selection and can be ignored (CHARLESWORTH and LANGLEY 1991). After formation of the new offspring population, transposition and excision are assumed to occur randomly throughout each genome under a Poisson process, with the TE copy number increasing with probability u per element and decreasing with probability v per element.

3.2.3 Model Parameters

Simulations were generally run for 10^6 generations, although in cases where non-equilibrium TE accumulation was especially rapid, simulations were run for fewer generations. At every 1000th generation, we surveyed the TE copy number on each chromosome and the number of sites where a TE was fixed across the entire population. At least 10 simulations were performed for each combination of parameters investigated. We set the parameters in our model to match estimates from *Drosophila* populations, where average transposition rates for all classes of elements are ~10^{-4} per element per generation, with excision rates at least one magnitude smaller (NUZHDIN and MACKAY 1995; SUH *et al.* 1995; VIEIRA and BIÉMONT 1997; PASYUKOVA *et al.* 1998; MASIDE *et al.* 2000). Since equilibrium requires equality of rates of origination and removal of elements, the strength of selection against segregating elements should be of the order 10^{-5}–10^{-4} per copy (CHARLESWORTH *et al.* 1994). Because of synergism between elements, the strength of selection depends on both a and b, so we set $a = 10^{-5}$ and let b vary from 10^{-6}–10^{-5}.

However, in order to achieve more simulations in a given amount of computation time, we adjusted the parameters following the scaling employed and validated in the simulations of DOLGIN and CHARLESWORTH (2006), which has also been used elsewhere (LE ROUZIC *et al.* 2007a; SÖDERBERG and BERG 2007). This scaling involves increasing the transposition–excision and selection parameters (u, v, a and b) by two orders of magnitude, while decreasing N and the number of generations by the same factor. Since we are maintaining the products of the deterministic forces and the effective population size constant, this should lead to the same evolutionary outcome, provided that all evolutionary forces are weak (EWENS 2004, Chap. 4). In the results below, we present the uncorrected parameter values as

used in the simulations, although to reflect biologically relevant values, all the parameters would need to be scaled 100-fold.

3.3 Results

With the quadratic fitness formula (CHARLESWORTH 1990; DOLGIN and CHARLESWORTH 2006), the expected equilibrium mean TE copy number under free recombination with low frequencies of elements at each occupied site is given by:

$$\bar{n} \approx \frac{u - a - v}{b}. \tag{3.2}$$

With the standard rate of crossing over across our entire simulated genome, including the small ("fourth") chromosome, we found that the TE copy number in simulations of large populations ($N = 10^3$) equilibrated rapidly and matched this predicted amount within ±1 element for all parameter combinations tested (results not shown). In the absence of excision, when we set the recombination rate on the fourth chromosome equal to zero, equilibrium copy numbers as predicted by Equation (3.2) were still established when the synergism coefficient, b, was large, implying strong synergistic interactions between insertions. Under these conditions, around 10% of the total number of elements in the genome were on the non-recombining fourth chromosome, which is directly proportional to the size of the chromosome in our simulations. However, with lower values of b, the genome accumulated more elements than the equilibrium expectations (Figure 3.1). This is the result of a build-up of elements on the non-recombining fourth chromosome, as TEs become fixed due to Muller's ratchet (MULLER 1964; FELSENSTEIN 1974). The shift between TEs either equilibrating or accumulating on the non-recombining chromosome occurs quite suddenly following slight changes in the synergism coefficient, b. By analyzing a single simulation run, we can observe this shift and see these fixations as discrete jumps in the TE copy number on the fourth chromosome between periods when the mean TE copy number remains fairly constant (Figure

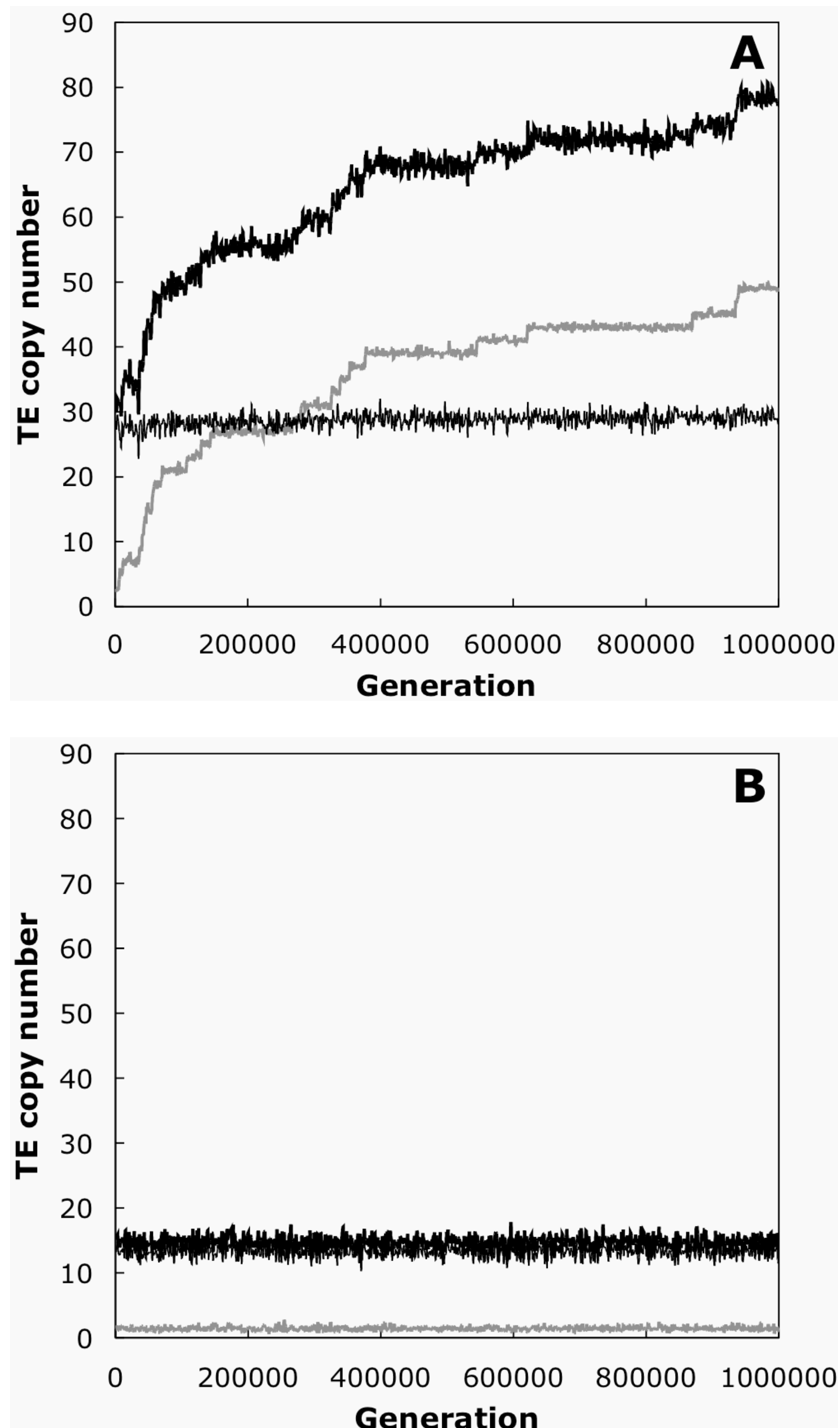


Figure 3.1: An example simulation run with weak (A) and strong (B) synergism when the fourth chromosome is completely non-recombining. Diploid mean TE copy numbers are shown for the entire genome (thick black line), the non-recombining fourth chromosome (thick grey line), and the three freely recombining chromosomes (thin black line) for simulation parameters N = 1000, u = 0.01, v = 0, a = 0.001, and b = 0.0003 (panel A) or b = 0.0006 (panel B).

3.1). This is in contrast to the freely recombining chromosomes, which maintained stable equilibrium copy numbers throughout the simulations, and the situation with strong synergism between elements, where no fixations were observed (Figure 3.1).

Classical mutation models usually assume a uniform genomic mutation rate and a constant selection coefficient for new mutations, resulting in a constant expected speed of Muller's ratchet (e.g. GORDO and CHARLESWORTH 2000a, 2001; STEPHAN and KIM 2002). On the other hand, population genetic models of transposable elements generally assume a fixed rate of transposition per element (LE ROUZIC and DECELIERE 2005), leading to greater genomic transposition rates with higher TE copy numbers. To counter this potential proliferation, the fitness effects of each insertion must be a decreasing function of TE abundance (CHARLESWORTH and CHARLESWORTH 1983). Thus, with more elements, two antagonistic forces can act: the ratchet may either be accelerated as a result of the increased genomic transposition rate, or decelerated owing to stronger selection against insertions.

Under the quadratic fitness equation used here, we found that the time between successive fixations generally increased as copy number increased, indicating that Muller's ratchet slows down as more elements accumulate on the non-recombining chromosome, and suggesting that the effects of stronger selection outweigh the higher genomic transposition rate (Figure 3.2). For example, for the two sets of parameters investigated in Figure 3.2, we performed a least-squares linear regression of time between fixations and "click" number, and found significantly positive slopes (Panel A: $F_{1,465} = 56.3$, $P < 0.0001$; Panel B: $F_{1,209} = 7.42$, $P = 0.007$). Our analysis is different from GORDO and CHARLESWORTH'S (2001) model of Muller's ratchet and transposable elements, since they essentially modified a mutational model to fit estimates of constant genomic transposition rates and assumed multiplicative fitness effects. However, it is consistent with an asexual mutational model that found that Muller's ratchet can effectively be halted if deleterious mutations have synergistic fitness effects (KONDRASHOV 1994; but see BUTCHER 1995), although in our model we never observed the fourth chromosome to reach a stable equilibrium copy number in simulations where TE fixations occurred, even over millions of generations (which equates to hundreds of millions of generations after scaling the parameters appropriately).

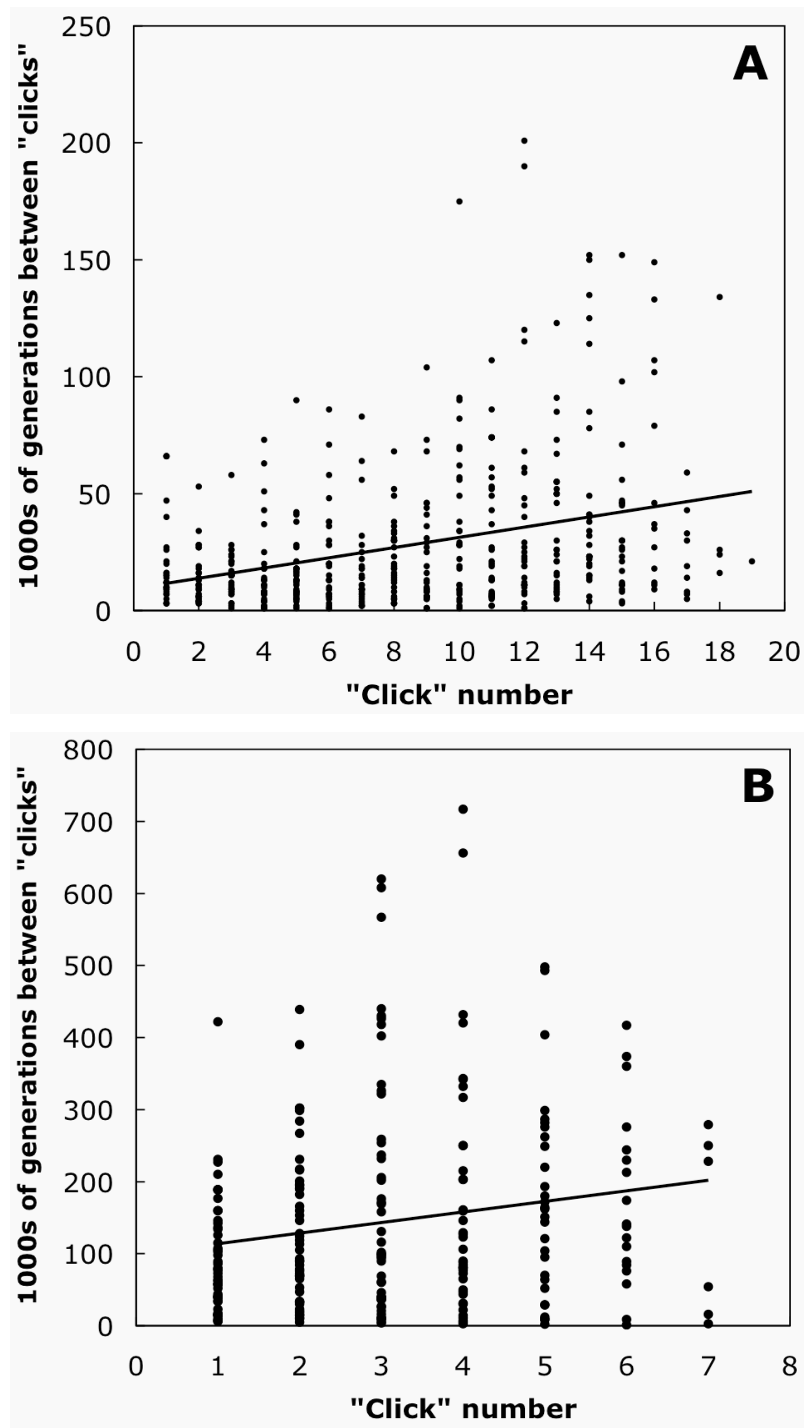


Figure 3.2: The time between "clicks" of the ratchet as a function of the ratchet number for simulations with a completely non-recombining fourth chromosome for simulation parameters: N = 1000, u = 0.01, v = 0, a = 0.001, and b = 0.0003 (panel A) or b = 0.0004 (panel B). Panel A shows the results of 30 simulations run for 5×10^5 generations, and panel B shows the results of 50 simulations run for 10^6 generations. The solid line shows the least-squared linear regression.

Table 3.1: The speed of Muller's ratchet on a null- or low-recombining chromosome. The average number of "clicks" per 10^5 generations is shown, assuming a constant rate of element fixation, with the recombination rate on the fourth chromosome either 0 (null recombination) or 10^{-5} between insertion sites (low recombination). The remaining parameters are: $u = 0.01$; $v = 0$; $a = 0.001$.

	Null recombination				Low recombination			
	***N* = 1000**		***N* = 500**		***N* = 1000**		***N* = 500**	
b	**"Clicks"**	**Runs**	**"Clicks"**	**Runs**	**"Clicks"**	**Runs**	**"Clicks"**	**Runs**
0.0002	39.4	20 *		^	12.9	10 **		^
0.0003	3.1	30 ***		^	0.50	10	9.95	10 **
0.0004	0.49	50	7.4	20 **	0.03	10	1.34	10
0.0005	0.12	30	1.86	10 ***	0	10	0.55	10
0.0006	0.01	30	0.62	10			0.22	10
0.0007	0	10	0.34	10			0.13	10
0.0008			0.15	10			0.02	10
0.0009			0.17	10				

All simulations are run for 10^6 generations, except where indicated:
* 10^5 generations; ** 2×10^5 generations; *** 5×10^5 generations.
^ denotes simulations where element fixation was so rapid that the computing power was insufficient to run the simulation for 10^5 generations.

If we ignore the fact that the rate of Muller's ratchet is not constant for the sake of comparison, and simply calculate a mean number of fixations per generation across the entire length of the simulation run, we have a rough measure with which to compare the speed with which TEs accumulate under different parameter conditions. Table 3.1 shows this average rate of fixation on the fourth chromosome under a range of synergism coefficients for two population sizes. As in Figure 3.1, it can be seen that element fixation was much more prevalent when synergism was weak (i.e. *b* was small). For a given strength of synergism, element fixation was also more likely to occur when populations were smaller, and fixations tended to be more rapid in smaller populations. We also considered the situation when local recombination was reduced but not completely absent. When the recombination frequency on the fourth chromosome was 10-fold lower than the rest of genome ($r = 10^{-4}$ between any two insertion sites, resulting in one crossover event per fourth chromosome every ~30 generations on average), no fixations were ever observed for any set of parameters tested (results not shown). With a 100-fold lower recombination frequency on the fourth chromosome ($r = 10^{-5}$ between any two insertion sites, resulting in one crossover event per fourth chromosome every ~300

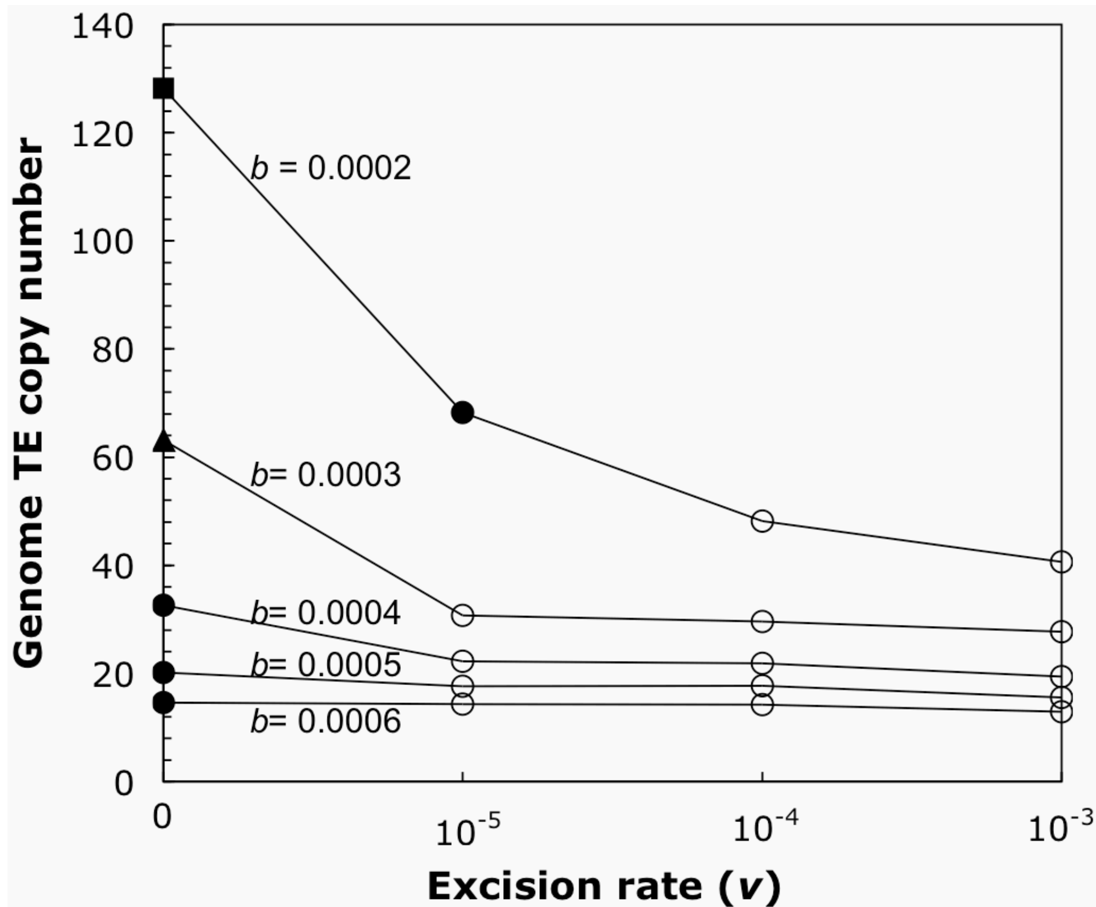


Figure 3.3: The effect of excision on element abundance. The mean genome TE copy number at the end of the simulation is shown for simulations with a completely non-recombining fourth chromosome for a range of synergism coefficients and excision rates, and the remaining simulation parameters were held constant: N = 1000, u = 0.01, a = 0.001. Simulations were run for 10^6 generations (denoted by circles), except for the triangle (5×10^5 generations) and the square (10^5 generations). Open circles indicate simulations that reached stable transposition–excision–selection equilibria, and closed symbols indicate simulations where fixations leading to non-equilibrium dynamics were observed. All data points are based on at least 10 simulation runs.

generations on average), element accumulation was widespread, but occurred at a considerably slower rate than in the complete absence of recombination (Table 3.1).

Our model thus far has assumed no element excision; therefore, once an element becomes fixed, Muller's ratchet has clicked irreversibly. While this lack of excision may be true for some classes of retrotransposons, which show little or no excision, the cut-and-paste transposition mechanism of DNA-based elements often leads to frequent element loss (CRAIG *et al.* 2002; HUA-VAN *et al.* 2005). We incorporated excision into our model, and inspected the effects of excision rates 1–3 orders of magnitude below the rate of transposition for a genome with a completely non-recombining fourth chromosome. For most parameter combinations tested, the presence of element excision halted the process of Muller's ratchet (Figure 3.3). Although temporary fixations were observed, these elements eventually became excised, allowing the population to reach a stable equilibrium close to that predicted by Equation (3.2). Only when the rate of fixation was much greater than the rate of excision did Muller's ratchet outpace excision, leading to continual element accumulation on the fourth chromosome (i.e. with b = 0.0002 and $v = 10^{-5}$, see Figure 3.3).

3.4 Discussion

Many authors have suggested that the reduction in the pressure of selection against TEs when recombination is reduced can lead to the build-up of elements in low recombination regions (e.g. CHARLESWORTH and LANGLEY 1991; DURET *et al.* 2000; BARTOLOMÉ *et al.* 2002; RIZZON *et al.* 2002; FONTANILLAS *et al.* 2007); here we provide the first theoretical analysis of this situation. We have shown that accumulation of TEs under the deleterious insertion model is expected to occur only in the genomes of small populations where recombination is severely reduced, when excision is extremely rare or absent, and with weak synergism between elements. These criteria should be empirically testable, yet might only be satisfied under a very limited range of biological circumstances. This is in contrast to the ectopic exchange model, which broadly predicts that the number of TE insertions is inversely related to

the recombination rate, even in very large populations (LANGLEY *et al.* 1988; MONTGOMERY *et al.* 1991; CHARLESWORTH *et al.* 1992b). Therefore, if TE accumulation is observed in low recombination regions, but the conditions necessary for build-up under the deleterious insertion model are not met, we can infer that the ectopic exchange model is implicated. Our model thus provides a first step in distinguishing between the leading models of TE distribution and abundance in regions of different recombination.

The effects of recombination on the deleterious insertion model were previously only investigated without any variation in recombination rate across the genome. CHARLESWORTH and CHARLESWORTH (1983) identified only minor effects of the recombination rate on TE abundance, and DOLGIN and CHARLESWORTH (2006) found large effects of the complete absence of recombination with asexuality, but no theoretical study has considered chromosomal regions with different recombination rates. This is important, because differences in the local recombination rate can drastically influence the effective population size experienced by insertions, as a result of the Hill-Robertson effects (HILL and ROBERTSON 1966; FELSENSTEIN 1974). The importance of effective population size for the abundance of TEs in general was first noted by CHARLESWORTH and CHARLESWORTH (1983), who found that smaller populations led to slightly greater equilibrium TE copy numbers in the presence of recombination. Later studies showed that small population size with sexual reproduction can often lead to unlimited increases in copy number, especially with high selfing rates, but the population sizes involved in such situations were unrealistically small for most natural populations (BROOKFIELD and BADGE 1997; WRIGHT and SCHOEN 1999).

Here we have used large, randomly mating populations that may be more representative of natural populations of organisms like *Drosophila*, focusing on TEs with weak fitness effects that can segregate in non-coding regions. We found that Hill-Robertson effects in low recombination regions can lead to stochastic TE accumulation. Even so, the population sizes used in this study are still probably smaller than those found in nature for *Drosophila* populations. Most of our simulations involved a population size of $N = 10^3$, which reflects a size of 10^5 following the scaling employed here (see Methods). This is an order of magnitude

lower than the effective population size of 10^6 typically assumed for *D. melanogaster* (KREITMAN 1983). Simulations run with larger populations indicated that TE accumulation was less likely and/or occurred at a slower rate (results not shown), but these simulations were too slow and computationally intensive to replicate many times. The possibility of TE accumulation in natural populations could thus be more limited than the already restrictive set of circumstances identified in our model, where larger population sizes reduce Hill-Robertson effects. On the other hand, our model did not consider other factors that might reduce the effective population size and potentially accelerate the rate of ratchet, such as population subdivision (COMBADÃO *et al.* 2007) or effects of background selection and selective sweeps at linked coding or regulatory sites (GORDO and CHARLESWORTH 2001). The latter effect could well be very important for non-recombining genomic regions like the fourth chromosome (SHELDAHL *et al.* 2003) and the neo-*Y* chromosome of *D. miranda* (BACHTROG 2003), where DNA sequence diversity is lower than the genome-wide average by more than a factor of ten, implying a large reduction in effective population size.

Another problem in relating our results to *Drosophila* data is that the non-recombining "fourth" chromosome in our model represented 10% of the genome. In reality, however, the fourth chromosome is much smaller. When we adjusted the chromosome sizes and made the fourth chromosome occupy only 1% of the genome, TE accumulation was never observed under any of the parameter conditions permitting accumulation on the larger chromosome that we used (results not shown). This suggests that selection is more effective at preventing TE build-up when the insertion rate of new elements into non-recombining regions is lower, consistent with general Muller's ratchet theory that fixations occur more rapidly with greater mutation rates (GORDO and CHARLESWORTH 2000a). Although we exaggerated the size of the fourth chromosome, a larger chromosome might more accurately approximate other non-recombining chromosomes including sex chromosomes, such as the neo-Y chromosome in *D. miranda*, which is riddled with TEs (STEINEMANN and STEINEMANN 1998; BACHTROG 2003). Our model also only considered genomic element copy numbers in the tens to hundreds, but the total number of active elements in the *Drosophila* genome is probably about an order of magnitude greater

(BARTOLOMÉ *et al.* 2002; KAMINKER et al. 2002; BERGMAN *et al.* 2006). However, given the inflated size of the fourth chromosome in our simulations, the reduced genomic copy number counterbalances the enlarged chromosome size, and makes the rate of new insertions into the fourth chromosome reasonably realistic.

The results of GORDO *et al.* (2002) show that, when Muller's ratchet operates, the effect of deleterious mutations on neutral diversity in non-recombining genomic regions is less than that predicted by equilibrium background selection theory. We have argued that the ratchet is unlikely to be causing a build-up of TEs on the fourth chromosome of *D. melanogaster*. It follows that their effect on neutral diversity should be adequately predicted by the background selection formula if TEs are contained by selection against the deleterious effects of element insertions on fitness (CHARLESWORTH 1996). But the predicted reduction in diversity due to background selection effects from the observed number of segregating TEs on the fourth chromosome is at least a factor of 1000-fold (CHARLESWORTH 1996; CARR *et al.* 2001), compared with an observed autosome to fourth chromosome silent diversity ratio of about 20 for *D. melanogaster* (SHELDAHL *et al.* 2003). This suggests that, as noted by CHARLESWORTH (1996), the background selection model greatly overpredicts the reduction in diversity. One possibility is that element insertions outside coding sequences are largely neutral on the fourth chromosome, due to the absence of meiotic ectopic exchange. These elements would not, therefore, exert background selection effects. This faces the difficulty, however, that retrotranspons do not show evidence for accumulation on the fourth chromosome (BARTOLOMÉ *et al.* 2002; KAMINKER *et al.* 2002; BERGMAN *et al.* 2006), as we would expect with this hypothesis. The solution to this dilemma may be that many members of this element class appear to have invaded *D. melanogaster* recently, and have not yet had time to accumulate on the fourth chromosome (SÁNCHEZ-GARCIA *et al.* 2005; BERGMAN and BENSASSON 2007).

Nonetheless, analyses of the *D. melanogaster* genome have revealed that many families of TEs, including both DNA-based elements and retrotransposons, tend to accumulate in euchromatic and heterochromatic regions where recombination is reduced (BARTOLOMÉ *et al.* 2002; KAMINKER *et al.* 2002; RIZZON *et al.* 2002; BERGMAN *et al.* 2006; FONTANILLAS *et al.* 2007), but the nature of the mechanisms

involved is obscure. The results of our model indicate that TE accumulation under the deleterious insertion model is only expected when element excision is absent or extremely rare. Since DNA transposons can excise from the genome, albeit at a low rate (NUZHDIN and MACKAY 1995; SUH *et al.* 1995; VIEIRA and BIÉMONT 1997; PASYUKOVA *et al.* 1998; MASIDE *et al.* 2000), Muller's ratchet is likely to be ineffective for this class of elements. It thus seems more likely that the ectopic exchange model explains the *Drosophila* data, because the non-recombining fourth chromosome is enriched in DNA-based elements with the ability to excise (BARTOLOMÉ *et al.* 2002; KAMINKER *et al.* 2002; BERGMAN *et al.* 2006). This is further supported by the observation that the density of DNA transposons parallels the recombination rate along the chromosomes (RIZZON *et al.* 2002), whereas our model shows that low to moderate amounts of recombination effectively prevent elements from accumulating.

Unlike DNA-based elements, retrotransposons rarely excise from the genome (CRAIG *et al.* 2002; HUA-VAN *et al.* 2005), and tend to be fairly homogenously distributed, except for accumulation in some regions with extremely low recombination rates (RIZZON *et al.* 2002; but see BARTOLOMÉ *et al.* 2002). Recent analyses have demonstrated that TEs are often fixed within populations in non-recombining portions of the *Drosophila* genome (BACHTROG 2003; BARTOLOMÉ and MASIDE 2004; MASIDE *et al.* 2005). BARTOLOMÉ and MASIDE (2004) used a PCR-based approach to determine the population frequency of a number of TE insertions in regions of suppressed or free recombination. They found that elements in regions of high recombination segregated at low frequencies and tended to be young, whereas null recombination regions were enriched in old elements at high frequencies or fixation, including several retrotransposons. This is at least qualitatively consistent with a model of retrotransposons accumulating and becoming fixed in low recombination regions due to Muller's ratchet. However, this model would also predict accumulation of retrotransposons on the fourth chromosome; as discussed above, this is not the case. The presence of fixed TEs is nonetheless compatible with the ectopic exchange model, because neutral elements can drift to high frequencies and fixation in low recombination regions, given that these regions have small effective population sizes.

DNA-based elements and retrotransposons show many differences in their respective intragenomic dynamics and distributions, and the causes of these differences remains unclear (HUA-VAN *et al.* 2005). Here we have shown that a selective mechanism that has often been suggested to explain the prevalence of TEs in regions of suppressed recombination can probably be ruled out for DNA-based elements, since they exhibit appreciable rates of element excision. The limited population survey data for retrotransposons in *Drosophila* is consistent with our model, but screening of more insertions in multiple populations will be needed to test whether fixations are in fact more prevalent in low recombination regions. With the sequencing of more *D. melanogaster* strains, we should be able to better understand the population level processes underlying our genomic observations. Although we have identified some necessary conditions for TEs to build-up through Muller's ratchet, determining conclusively which selective mechanisms are involved in causing TEs to accumulate in low recombination regions of the genome requires more experimental and theoretical analyses.

4 Inbreeding and Outbreeding Depression in *Caenorhabditis*

The work presented in this chapter was published in**:**

- DOLGIN, E. S., B. CHARLESWORTH, S. E. BAIRD, and A. D. CUTTER, 2007 Inbreeding and outbreeding depression in *Caenorhabditis* nematodes. *Evolution* **61**: 1339-1352.

Contributing authors:

- I carried out the laboratory tests, performed the analysis, and wrote the manuscript.
- B. CHARLESWORTH advised on the project, and helped write the manuscript.
- S. BAIRD collected and provided the *C. remanei* isolates.
- A. CUTTER advised on the lab work and methodology, and helped write the manuscript.

4.1 Introduction

Mating between close relatives causes an increase in homozygosity, which usually results in a decline in fitness known as inbreeding depression. Inbreeding depression is a central factor in the evolution of mating systems, particularly in relation to the evolution of the rate of self-fertilization in hermaphroditic organisms (JARNE and CHARLESWORTH 1993; UYENOYAMA *et al.* 1993). Some models that incorporate inbreeding depression predict the evolution of either complete selfing or complete outcrossing (e.g. LANDE and SCHEMSKE 1985). Nevertheless, mixed mating

systems abound, and other theoretical studies have found conditions under which partial selfing is evolutionarily stable (see review by GOODWILLIE *et al*. 2005).

The nematode *Caenorhabditis elegans* exhibits a mating system in which hermaphrodites are self-fertile, but can outcross only with males. This androdioecious reproductive mode makes *C. elegans* a useful system for studying the evolution of outcrossing versus selfing. Under both laboratory and natural conditions, however, males are rare (HODGKIN and DONIACH 1997; BARRIÈRE and FÉLIX 2005, 2007). Studies of molecular variation in wild-caught strains of *C. elegans* indicate that selfing is the primary mode of reproduction, with low but detectable rates of outcrossing (DENVER *et al*. 2003; BARRIÈRE and FÉLIX 2005, 2007; SIVASUNDAR and HEY 2005; CUTTER 2006).

There is an ongoing debate about the evolutionary significance of males in *C. elegans*. Although males are present only at extremely low population frequencies, a significant portion of the genome is dedicated to male-specific functions (JIANG *et al*. 2001), which appears to be maintained by selection (CUTTER and WARD 2005). One obvious selective advantage to males is their escape from any inbreeding depression caused by the selfing of hermaphrodites, which is thought to be important in other androdioecious species (LLOYD 1975; CHARLESWORTH 1984; OTTO *et al*. 1993; RIESEBERG *et al*. 1993; WEEKS *et al*. 1999; STEWART and PHILLIPS 2002; CUTTER *et al*. 2003). However, inbreeding depression has not been observed in *C. elegans* for life span (JOHNSON and WOOD 1982; JOHNSON and HUTCHINSON 1993), brood size (CHASNOV and CHOW 2002), and various other life history traits (JOHNSON and HUTCHINSON 1993).

This is consistent with theoretical analyses, which show that very high levels of inbreeding are associated with large reductions in the frequencies of deleterious recessive or partially recessive mutations, leading to reduced inbreeding depression (CHARLESWORTH and CHARLESWORTH 1998). The restriction of gene flow by a predominance of selfing may also lead to greater levels of population subdivision, which could further reduce inbreeding depression (WALLER 1993; THEODOROU and COUVET 2002; WHITLOCK 2002; GLÉMIN *et al*. 2003). In addition, restricted recombination, resulting from low rates of outcrossing and migration (BARRIÈRE and FÉLIX 2005, 2007; CUTTER 2006), could lead to the accumulation of different

favourable combinations of alleles in different local populations, resulting in hybrid breakdown (outbreeding depression) following outcrossing (TEMPLETON 1986).

Phylogenetic evidence indicates that the ancestor of *C. elegans* was gonochoristic (separate males and females)—changes from gonochorism to hermaphroditism have independently occurred at least 10 times in rhabditid nematodes (KIONTKE *et al*. 2004; KIONTKE and FITCH 2005). In order to understand the evolution and regulation of mating systems, comparative studies of inbreeding depression in conspecific and congeneric populations with different modes of reproduction have been undertaken in other groups (e.g. HOLTSFORD and ELLSTRAND 1990; DEMEESTER 1993; JOHNSTON and SCHOEN 1996). Given the prominence of *Caenorhabditis* species as model organisms in biological research, it is surprising that little is known about inbreeding depression levels in the genus. Even in *C. elegans*, all the studies mentioned above involved the N2 strain, which has been maintained in the laboratory for thousands of generations. Natural isolates that have not adapted to laboratory conditions might exhibit different levels of inbreeding depression (STEWART and PHILLIPS 2002), especially in view of the observed variation among strains in male production and persistence within laboratory populations (TEOTÓNIO *et al*. 2006).

The purpose of this study was to examine the relationship between breeding system and inbreeding depression, using two *Caenorhabditis* species with contrasting mating systems, *C. elegans* and *C. remanei*. To this end, we used worms recently derived from the wild, and assayed fitness-related traits using similar methodologies for the two species, in order to obtain estimates of the levels of inbreeding depression. We found that the outcrossing species, *C. remanei*, suffered strong multi-generational inbreeding depression, with the majority of inbred lines going extinct. In contrast, the selfing species, *C. elegans*, mostly exhibited outbreeding depression. We discuss the implications of these results for patterns of genetic diversity and the evolution of mating systems.

4.1 Materials and Methods

4.1.1 Nematode Populations

C. elegans strains were recently isolated from three localities in France: Franconville (48°98N, 2°23E), Hermanville (49°28N, 0°32W) and Merlet (44°45N, 4°42E) (BARRIÈRE and FÉLIX 2005). These strains were selfed for a few generations to reduce any within-strain heterozygosity before freezing, and were only thawed shortly before initiating the present experiment. Five crosses were established, which represent a variety of within and between population comparisons: three crosses between strains from different populations (JU318 × JU370; JU322 × JU399; JU342 × JU466), one cross between strains isolated within 10 centimeters of each other in Franconville (JU364 × JU368), and one cross between strains isolated 15 meters apart in Merlet (JU314 × JU323). All crossed strains were selected on the criterion of being divergent at a number of amplified fragment length polymorphism (AFLP) markers found among a larger data set, as described in BARRIÈRE and FÉLIX (2005), so as to ensure the crosses were between genetically distinct strains (Table 4.1). We also present two other measures of genetic divergence from two subsequent studies that used these strains, the number of silent-site pairwise single nucleotide polymorphisms (SNPs; CUTTER 2006) and the mean squared difference in microsatellite repeat length (BARRIÈRE and FÉLIX 2007), and test whether genetic distance affects the magnitude of the observed effect.

Males of *C. elegans* were generated in each strain by heat-shock at 26–28°C for ~6 hours, and the worms were then returned to 20°C to allow selfing. The resulting male offspring were then crossed in abundance to hermaphrodites, in order to maintain populations with a ~50:50 sex ratio. Each line was then maintained at 20°C for three generations as both a mixed-sex and a pure hermaphrodite population.

Crosses were carried out in four separate experimental blocks at different time points, with 2–3 genetically distinct crosses per block, and each particular cross was performed twice. It is of interest to note that another cross was also attempted, JU393 × JU407, but the male mating performance of JU393 was so poor that mating frequency was extremely low and these data were discarded. The particularly poor

Table 4.1: Strains used in crosses, with geographic source in parentheses, and molecular divergence data. The number of differing AFLP markers is from a total of 149 measured and 31 found polymorphic. The number of silent site SNPs is from six regions on two chromosomes spanning a total of 3372.4 silent sites across all loci, with the diversity measure π_{si} shown in parentheses. The mean squared difference in microsatellite repeat length ($\delta\mu^2$) is from six microsatellite loci across five chromosomes.

Strain 1	Strain 2	Geographic distance	Number of differing AFLP markers†	Number of silent-site pairwise SNPs (π_{si})*	Mean squared difference in microsatellite repeat length‡
JU364 (Franconville)	JU368 (Franconville)	< 10 cm	9	0 (0.0)	0
JU314 (Merlet1)	JU323 (Merlet2)	~ 15 m	14	21 (0.00614)	28.3
JU318 (Merlet1)	JU370 (Franconville)	570 km	15	5 (0.00144)	3.5
JU342 (Merlet2)	JU366 (Franconville)	570 km	8	21 (0.00610)	21.7
JU322 (Merlet2)	JU399 (Hermanville)	655 km	9	17 (0.00494)	17.8

† BARRIÈRE and FÉLIX 2005; * CUTTER 2006; ‡ BARRIÈRE and FÉLIX 2007.

male mating ability of the JU393 strain was also independently observed by H. TEOTÓNIO (personal communication).

C. remanei populations were established from dauer larvae found living in association with terrestrial isopods in the Wright State University (WSU) Biological Preserve in Southwestern Ohio, USA (39°47N, 84°03W). Sampling was done following the protocol of BAIRD (1999). Briefly, *Trachilipus rathkii* isopods were collected during a single day from several locations within a 0.1 km radius in a wooded area of the WSU Biological Preserve, and sacrificed on 60-mm NGM-lite agar plates seeded with *Escherischia coli* OP50. Nematodes obtained in these collections were considered to be from a single genetically diverse, random mating population, consistent with current knowledge of *C. remanei* from this area (CUTTER *et al*. 2006a).

Plates were periodically monitored for ~2 days for the appearance of nematodes, and any worms obtained were transferred to a different plate. In total, 30 females and 24 males were isolated, allowed to develop to adulthood, and to mate randomly for ~1 day, allowing the possibility of multiple matings, in order to initiate the study population. The males were then killed and the 30 gravid adult females were subdivided onto three fresh plates, with 10 females per plate, to establish three subpopulations. These plates were then maintained as large outbred populations for ~6 generations before initiating the experiment, and concurrently frozen for storage at –80°C. Except where specifically noted, worms of both *C. elegans* and *C. remanei* were maintained at 20°C on 35mm NGM-lite agar Petri plates seeded with *E. coli* OP50, using standard techniques (SULSTON and HODGKIN 1988).

4.2.2 Experimental Crosses and Life-History Assays

In *C. elegans*, for each cross we measured fecundity and longevity in four classes of hermaphrodite: pure strain F1s from each strain, and the two reciprocal hybrid F1s. Individuals of the same class were not independent of each other, as they often came from the same family, considering multiple hermaphrodite offspring of a cross involving a single mother and 5–6 fathers to belong to the same family. Altogether, 767 worms were assayed, composed of an average of 38 worms per class in each cross, with a mean family size of 5.85, yielding a total number of 178,861 offspring. The general features of the protocol are presented in Figure 4.1.

We first synchronized populations and cleaned plates of any bacterial contamination by using alkaline hypochlorite (SULSTON and HODGKIN 1988). To ensure that the hermaphrodites assayed were the result of crossing rather than selfing, we placed 5–6 larval stage L4 males from the same strain with one L4 hermaphrodite. After 2–3 days of mating, each hermaphrodite was moved to a fresh plate to lay eggs for 5 hours. After 2–3 days, we verified that the F1 offspring that developed from these eggs had a male to hermaphrodite ratio that did not differ significantly from 1:1, using a χ^2-test with one degree of freedom. These F1 hermaphrodites were then used for fecundity and longevity assays.

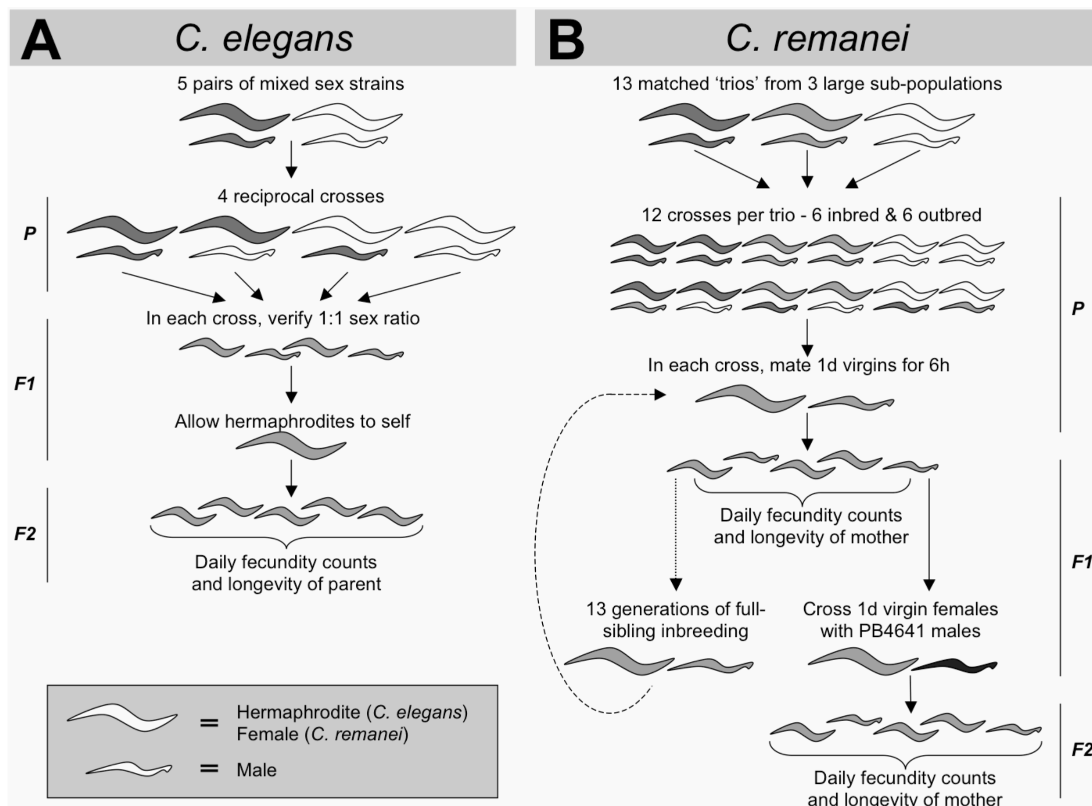


Figure 4.1: Experimental protocol for *C. elegans* (A) and *C. remanei* (B). The fitness-related traits were daily fecundity, giving a measure of total brood size and relative fitness (w), and longevity. The experimental protocol was designed to measure similar components of fitness for the different breeding systems.

In *C. remanei*, multiple plates from each subpopulation were set up with a single female and a single male at the L4 stage in order to obtain full sibling offspring. Only plates with many offspring (>300) were subsequently used, to guarantee healthy outbred starting populations. To ensure that the individuals being crossed were as unrelated as possible, we only used crosses between subpopulations for the outbred crosses. This is because individuals within subpopulations may have mated during the ~6 generations between the field collection and establishment of the subpopulations, and initiating the experiment. To this end, we randomly chose one plate from each of the three subpopulations to create 13 "trios".

Within each trio of subpopulations, we set up six inbred crosses, two from each subpopulation, and all six possible outbred crosses between each subpopulation,

including both the reciprocal crosses (Figure 4.1B). We then measured productivity and longevity in the resulting outbred and full-sibling inbred females, and in the F1 progeny of these individuals crossed to males of the standard strain, PB4641 (selected because it is a highly inbred strain, which is being used for genomic sequencing). Using PB4641 males allows us to standardize the male contribution in all treatments and isolate the genetic effects of inbreeding depression on the F1 females. A total of 215 worms were assayed, 71 inbred and 63 outbred individuals, of which 46 inbred F1 and 35 outbred F1 progeny were crossed with PB4641. A total of 60,750 offspring were counted.

Additionally, we maintained a number of inbred lines for 13 generations of full-sibling mating. We attempted to propagate 39 randomly selected lines, representing all three subpopulations taken from inbred F1 progeny, by placing full-sibling single males with single females, isolated as L4s. In establishing our inbred lines, we set up two mating plates of each inbred line in each generation, to allow for failure to mate. If no progeny were found on either plate, the line was deemed to have gone extinct. Otherwise, one of the plates with progeny was selected at random to propagate the inbred line. After 13 generations, we measured the fecundities of 6 classes of individuals from the extant inbred lines: (1) full-sibling inbred females, (2) inbred females crossbred with males from different inbred lines, (3) outbred females thawed from the three subpopulations frozen before initiation of the experiment, (4) inbred females crossed to PB4641 males, (5) crossbred F1 females whose parents were from different inbred lines crossed to PB4641 males, and (6) outbred females crossed to PB4641 males. At this stage, a further 170 worms were assayed and 38,106 offspring counted. We also froze the extant inbred lines at –80°C and repeated these productivity measures again in a second block.

Fecundity and longevity assay methods were similar between the two species, with a few notable exceptions. In *C. elegans,* L4 hermaphrodites were isolated onto individual plates (there was no discernible difference in the timing of development between pure-strain and hybrid worms). Once they had developed into adults, they were allowed to lay eggs, and were transferred to fresh plates every 24 h for 3 days, and then to new plates to measure late fecundity and longevity. In *C. remanei*, male and female virgins were isolated at the L4 stage and kept separate for 40–44 hours

until they had become ~1 day-old adults. Although serially inbred lines of *C. remanei* had delayed growth rates, we ensured that they had molted to adults many hours before the time of mating. Female reproduction in *C. remanei* is limited by the number of sperm transferred upon mating. Consequently, we placed a single male and female together on a fresh plate and permitted them to mate for 6 h to mirror the sperm-limitation experienced by selfing protandrous *C. elegans* hermaphrodites, where spermatogenesis precedes oogenesis, and once egg maturation begins, the hermaphrodite possesses its full complement of sperm. The male was then killed and the female moved to a new plate for 18 h, transferred every 24 h for 2 more days, and then moved to new plates to measure late fecundity and longevity. Plates without any offspring were observed in 13 of 156 crosses in the parental generation, with no difference between inbred and outbred crosses ($\chi_1^2 = 0.08$, $P = 0.78$). It was assumed that mating had not been successful on these plates, and they were omitted from further analyses.

For both species, after the timed plate transfers for days 1–3, the worms were transferred after 2 days, and subsequently only if offspring or bacterial contamination was observed, although these were both rare events. Worms were then checked every 1–2 days, and were deemed to have died if they did not respond to gentle agitation with a platinum pick. Eggs on the laying plates were allowed to hatch and develop for 3 days, and were then counted to give daily egg-to-late-larval fecundities for days 1–3. For *C. remanei*, day 1 was divided between the first 6 h mating period and the subsequent 18 h, which was denoted as day 1.5 for the calculation of *w* (see below). Late fecundity beyond day 3 was summed and included in measures of total brood size.

Daily fecundities were used to generate two fitness components: brood size (the unweighted sum of all progeny over the entire lifespan including late productivity) and relative fitness, *w*. This second measure is proportional to the expected fitness of an age-structured population, defined as $w = \sum_x e^{-rx} l_x m_x$ (CHARLESWORTH 1994), where $l_x m_x$ is the product of survivorship to age x and fecundity at day x, and r is a constant scaling term, equal to the growth rate of the population as a whole. Due to the short development time of these worms and the sperm-limited fecundity of *C. elegans* hermaphrodites or singly-mated *C. remanei*

females, the timing of egg-laying is the principal factor affecting *w*. For *C. elegans*, *r* was calculated by defining the grand mean fitness of the pure strains as $\overline{w}_{\text{pure-strain}} = 1$, obtained using $l_x m_x$ pooled across all pure strain worms. This *r*-value was used to weight the productivity by a negative exponential function of age. For *C. remanei*, *r* was calculated by defining the mean fitness of the outbred individuals in the parental cross as $\overline{w}_{P-\text{outbred}} = 1$. This value of *r* was then used to calculate *w* for the crosses with PB4641 males and for the crosses after 13 generations of inbreeding.

4.2.3 Statistical Analysis

Longevity was analyzed by Kruskal-Wallis tests, and brood size and *w* were analyzed by general linear mixed models, using the JMP statistical package, version 5.1 (SAS INSTITUTE, Cary, NC, USA). Residuals were normally distributed for brood size and *w* in *C. elegans*, and for *w* in *C. remanei*, but brood size data for *C. remanei* had to be square-root-transformed to normalize the residuals, because of the large differences between inbred and outbred worms. For *C. elegans*, fixed factors were breeding class (i.e., hybrid versus pure-strain), cross identity, block, block × breeding class, block × cross identity, and breeding class × cross identity. A random effect in the analysis was family, with Z-statistics used to test significance. Heterosis was also measured within each cross. Heterosis is usually defined as the relative increase in fitness of hybrids between strains due to increased heterozygosity. We measure it as the excess of the mean fitness of the hybrid F1s, relative to the mean of the pure-strain values (FALCONER and MACKAY 1996, p. 253), defined here as

$$H_j = \left[\frac{\overline{X}_j(\text{hybrid}) - \overline{X}_j(\text{pure strain})}{\overline{X}_j(\text{pure strain})} \right] \tag{4.1}$$

where $\overline{X}_j(\text{hybrid})$ and $\overline{X}_j(\text{pure strain})$ represent the means for the particular trait under investigation, *X*, of hybrid and pure-strain F1s within the *j*th cross. Thus, a positive value of *H* denotes inbreeding depression, and a negative value of *H* indicates outbreeding depression. The means within the same cross were compared using *t*-tests.

For *C. remanei*, inbreeding level and trios of matched subpopulations were included as fixed factors in the analysis of brood size and *w* for the assays of the first parental (P) generation of inbreeding and the F1 cross with PB4641 males. For the P generation cross, the inbreeding levels were sib-mated and randomly-mated, and for the F1 × PB4641 cross the inbreeding levels were inbred and outbred. For the assays after 13 generations of inbreeding, brood size and *w* were measured, but not longevity. Two general linear models were analyzed, with block and inbreeding class included as fixed factors. In the first analysis, the inbreeding classes were inbred (within lines), crossbred (the crosses between different inbred lines) and outbred, while in the second the F1s of these classes crossed to PB4641 males were considered. The means were compared between the fully inbred, crossbred and outbred females and the F1s crossed with PB4641 males, using Tukey tests. To analyze the rate of extinction of inbred lines over the course of 13 generations of serial inbreeding, we performed random permutation tests for a least-squares linear regression model, in order to determine appropriate significance thresholds that take into account the correlations among extinction rates between successive generations. The permutation tests were performed by holding the generation numbers constant and permuting the per generation rates of extinction 10^7 times, and then comparing the resulting r^2 values.

4.3 Results

4.3.1 Outbreeding Depression in *C. elegans*

For both total brood size and relative fitness (*w*), pure-strain F1s performed significantly better than hybrids ($F_{1,631} = 16.4$, $P < 0.0001$ for brood size; $F_{1,631} = 22.7$, $P < 0.0001$ for *w*). Both of these traits were affected by the cross identity ($F_{3,631} = 10.9$, $P < 0.0001$ for brood size; $F_{3,631} = 8.3$, $P < 0.0001$ for *w*), and the magnitude of outbreeding depression also varied significantly between different crosses (as indicated by breeding class × cross identity: $F_{4,631} = 8.5$, $P < 0.0001$ for brood size; $F_{4,631} = 8.3$, $P < 0.0001$ for *w*). The overall trend was significantly

towards outbreeding depression with mean heterosis values of –0.064 and –0.093 for brood size and *w*, respectively (Figure 4.2). When examining each cross independently, however, outbreeding depression was only significant in some crosses (3 out of 5 for brood size; 2 out of 5 for *w*), and one cross showed a significant positive heterosis value for brood size (Figure 4.2).

The block in which assays were carried out was also a significant factor for brood size ($F_{2,631}$ = 11.1, P < 0.0001), but not for *w* ($F_{4,631}$ = 2.81, P = 0.06). Interactions between block and genetic factors for brood size were not significant ($F_{3,626}$ = 0.64, P = 0.59 for block × cross identity; $F_{3,626}$ = 2.1, P = 0.13 for block × breeding class). This indicates that experimental noise was introduced into the data by undertaking the experiment over 4 blocks. However, the lack of significant interactions or strong block effects on *w* suggests that these effects were, for the most part, in the same direction for all genotypes.

Relative fitness*, w,* is measured from daily progeny counts and so is not independent of total brood size. These two measures were significantly positively correlated (r = 0.475; t_{764} = 13.1, P < 0.0001), with the magnitude and direction of outbreeding depression for both brood size and *w* in the same direction (r = 0.88), although this is limited by sample size (t_3 = 1.31, P = 0.14). Neither geographic proximity nor the number of polymorphic AFLP markers shared between the crossed strains affected the direction of outbreeding depression (P > 0.59 for all regression analyses). There was, however, a non-significant trend towards a positive relationship between the magnitude of outbreeding depression and both the number of pair-wise single-nucleotide polymorphism differences ($F_{1,3}$ = 4.49, P = 0.12 for brood size; $F_{1,3}$ = 2.06, P = 0.25 for *w*) and the mean squared difference in microsatellite repeat length ($F_{1,3}$ = 12.46, P = 0.03 for brood size; $F_{1,3}$ = 6.98, P = 0.07 for *w*) (see Table 4.1). This suggests that more divergent strains may exhibit stronger outbreeding depression. More crosses would be needed to test this possibility.

Unlike for total brood size and *w*, hybrids and pure strain worms showed no difference in longevity (χ^2_1 = 0.12, P = 0.73). Although the five crosses exhibited different life spans (χ^2_4 = 20.7, P = 0.0004), there was no difference in the longevity of pure-strain versus hybrid F1s in any of the individual crosses (all P > 0.05; see

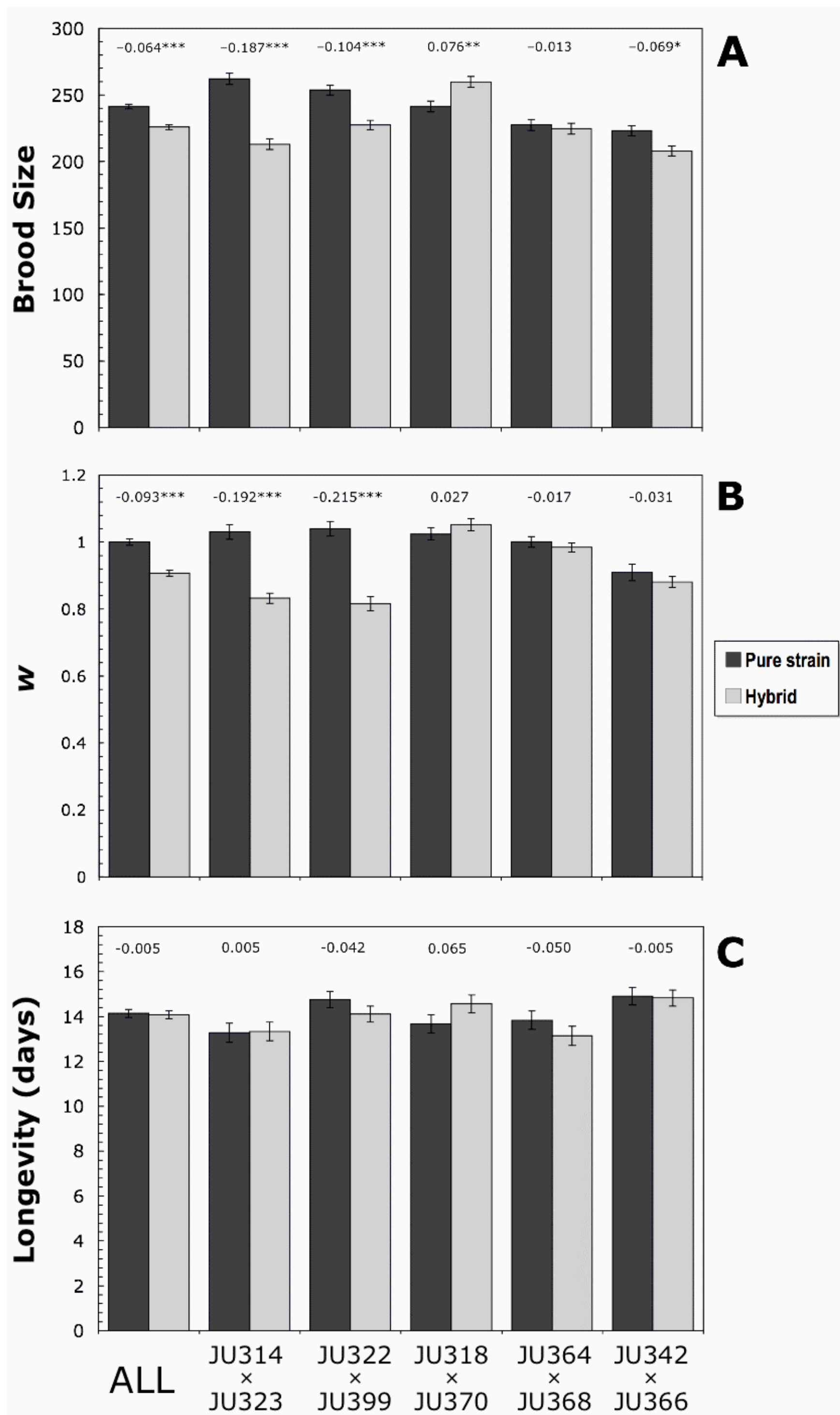


Figure 4.2: Measurements of mean pure-strain (darkly shaded bars) and hybrid (lightly shaded bars) brood sizes (A), relative fitnesses (B), and longevities (C) by cross identity and overall in *C. elegans*. Numbers above the bars show heterosis values, calculated as (hybrid/pure-strain) – 1. Asterisks indicate significance of difference between pure-strains and hybrids in a Student's *t*-test. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Error bars indicate ± 1 SE.

Figure 4.2C). Blocks also had significant effects on longevity (χ_3^2 = 35.1, $P < 0.0001$), although breeding class remained non-significant when considering any particular block independently (all $P > 0.05$). Family was a significant effect for brood size ($Z = 4.8$, $P < 0.0001$), *w* ($Z = 5.2$, $P < 0.0001$), and longevity ($Z = 1.9$, $P = 0.03$), suggesting a considerable amount of variation among parents of the same genotype. This is probably not due to genetic diversity in the parental strains, since the strains used here were selfed for a few generations before freezing, and heterozygosity was found to be quite low in a sub-sample of fresh isolates collected from the same locations as the ones being used here (BARRIÈRE and FÉLIX 2005). It is more likely that this is caused by variation in the ages of the mothers. Eggs laid by older worms tend to be more developed than those laid by younger worms. Therefore, even though we used a timed egg lay, the ages may not have been developmentally synchronized if the mothers were of different ages (PETERS *et al.* 2003).

4.3.2 Inbreeding Depression in *C. remanei*

Full-sibling mated females had a lower total brood size and *w* compared with randomly mated females, although this effect was not significant. ($F_{1,120}$ = 1.43, $P = 0.23$ for brood size; $F_{1,120}$ = 1.35, $P = 0.25$ for *w*). However, when these females' F1 progeny were crossed with PB4641 males, inbreeding level was a significant factor for both brood size ($F_{1,67}$ = 4.02, $P = 0.049$) and *w* ($F_{1,67}$ = 7.77, $P = 0.007$) (Figure 4.3). Our experimental design allows us to try to disentangle the sources of inbreeding depression. The poorer performance of sib-mated females in the parental generation, although non-significant (probably because of high variance in the productivity-related traits), can possibly be attributed to embryonic or larval inviability. The difference in the F1s crossed to PB4641 males indicates reduced egg production and/or a reduction in the timing of egg-laying in the inbred females, since the number of sperm and the mating performance of the PB4641 fathers are invariant. In both generations among crosses, inbreeding level had no effect on longevity (χ_1^2 = 0.01, $P = 0.91$ for parentals; χ_1^2 = 0.09, $P = 0.77$ for F1 × PB4641).

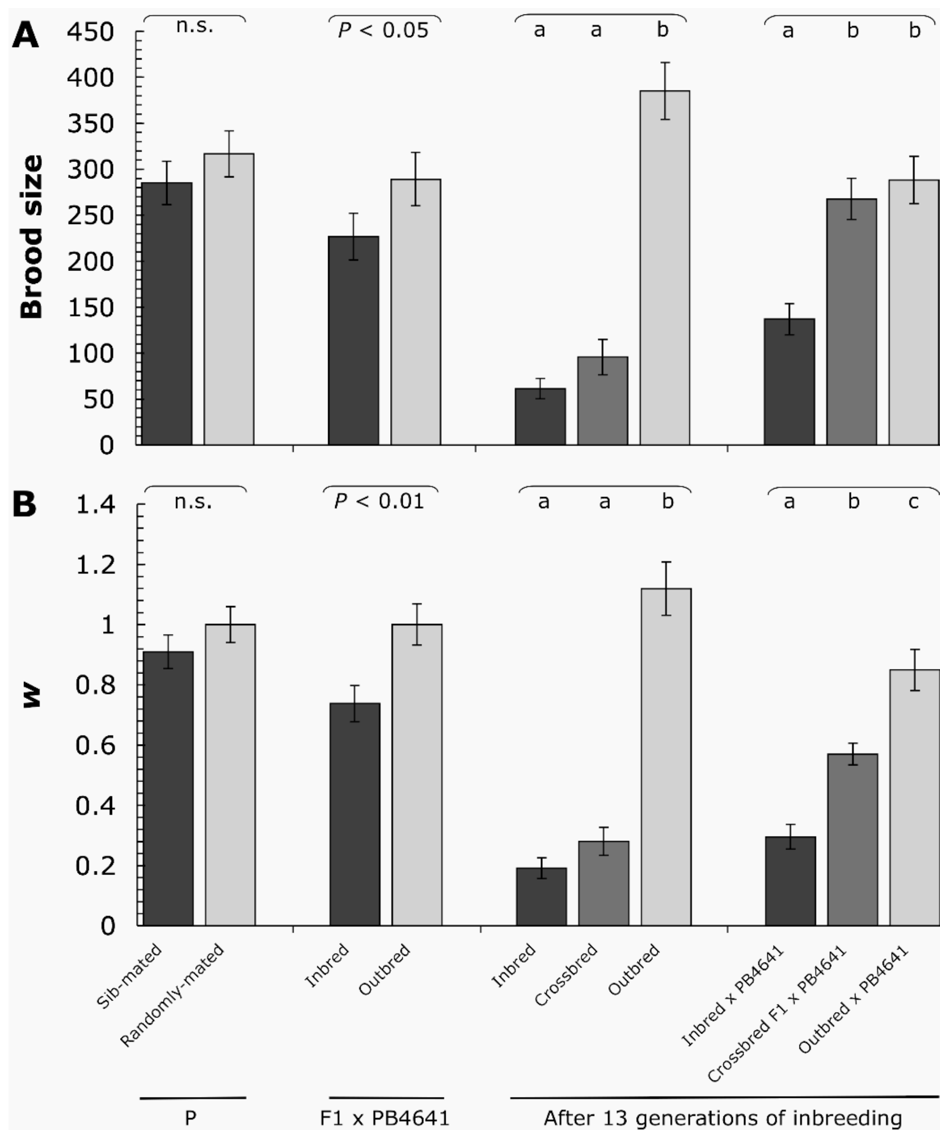


Figure 4.3: Measurements of brood size (A) and relative fitness (B) in *C. remanei* for the parental generation (P), the F1 females mated to PB4641 males (F1 × PB4641), and crosses after 13 generations of full-sibling mating. Darkly shaded bars denote sib-mated females and their inbred progeny, lightly shaded bars denote randomly mated females and their outbred progeny, and intermediate shaded bars denote inbred females crossbred with males from different inbred lines and their progeny. Post-hoc significance tests were done separately for each of the four sets of crosses – Student's *t*-tests for *P* and F1 × PB4641, and two separate Tukey tests for crosses after 13 generations of inbreeding. Each comparison is represented by a curved-line above the bars being evaluated with the significance denoted underneath. Identical letters within a comparison indicate non-significant differences, but letters across comparisons or across panels are unrelated. Error bars indicate ± 1 SE.

There were no significant effects of the trios of randomly matched sub-populations on any of the measures in either of the crosses.

For the assay of total brood size and *w* on the five extant lines after 13 generations of full-sib inbreeding, the effects of inbreeding level were highly significant ($F_{2,72}$ = 51.3, $P < 0.0001$ for brood size; $F_{2,72} = 56.4$, $P < 0.0001$ for *w*). Outbred females had greater brood size and relative fitness than inbred females mated to males from the same ("inbred") or different ("crossbred") inbred lines (Figure 4.3). Crossbred females had ~50% greater mean performance than inbred females, although this effect was not statistically significant due to large variances ($q_{72,\,3} = 1.67$, $P > 0.20$ for brood size; $q_{72,\,3} = 1.16$, $P > 0.50$ for *w*). The lower values for crossbred versus outbred females suggests that these inbred parents have reduced gamete production and/or mating ability, although the increased performance of crossbred versus inbred females suggests that larval inviability may play a role in the source of inbreeding depression as well.

Taking the F1 progeny females from the crosses between the strains surviving after 13 generations of full-sibling inbreeding and mating them to PB4641 males also yielded a highly significant effect of inbreeding level on fitness ($F_{2,92} = 12.3$, $P < 0.0001$ for brood size; $F_{2,92} = 29.5$, $P < 0.0001$ for *w*). Outbred females crossed to PB4641 males had greater trait values than inbred females crossed to PB4641 males (Figure 4.3), with the reduced performance of the inbred females being significantly greater after 13 generations of inbreeding as compared to after 1 generation ($t_{70} = 2.91$, $P = 0.005$ for brood size; $t_{70} = 5.95$, $P < 0.0001$ for *w*).

Crossbred F1 females mated with PB4641 males performed significantly better than inbred females ($q_{92,\,3} = 6.00$, $P < 0.001$ for brood size; $q_{92,\,3} = 6.21$, $P < 0.001$ for *w*); however, despite no difference in the mean total brood sizes for crossbred F1 females and outbred females mated to PB4641 males ($q_{92,\,3} = 1.10$, $P > 0.50$), *w* was significantly lower for crossbred F1 females ($q_{92,\,3} = 5.92$, $P < 0.001$). Because crossbred F1 individuals are the progeny of unrelated inbred lines, they have an inbreeding coefficient of 0. The near equivalence of brood sizes indicates that crossbred F1 females had fecundity levels restored to outbred levels, but the reduction in *w* suggests a maternal effect on the timing of egg production, since the parental lines of the crossbred F1 female had been inbred for 13 generations. There

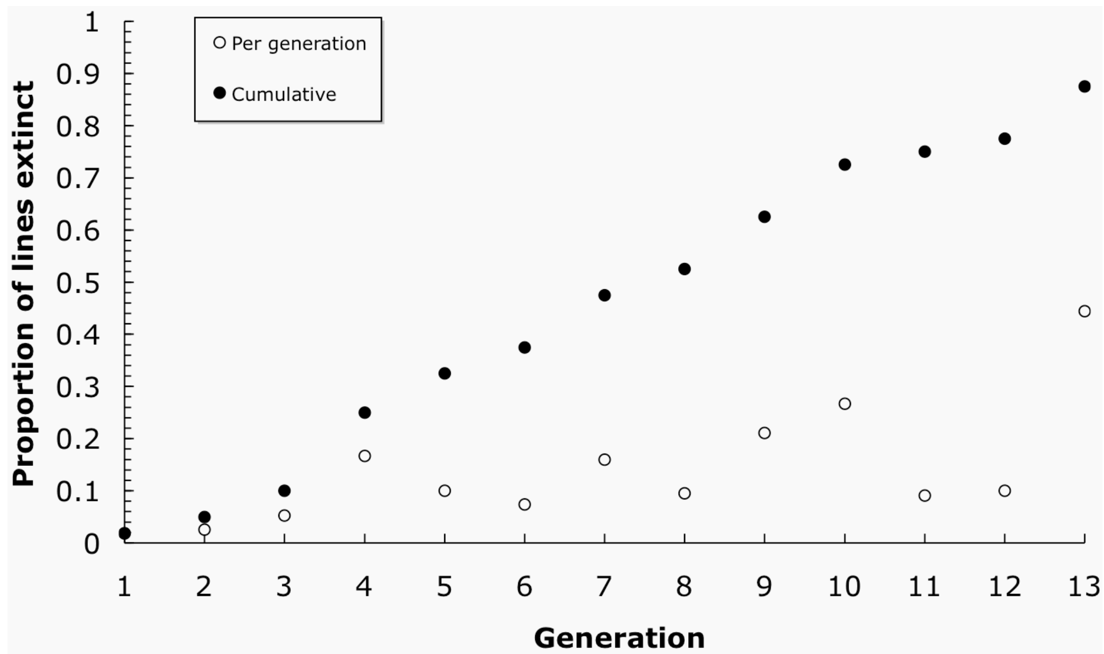


Figure 4.4: The rate of extinction of inbred lines in 13 generations of serial inbreeding. The open circles show the per-generation proportion of lines going extinct; the closed circles show the cumulative proportion.

was little evidence of purging of deleterious mutations during the process of inbreeding, since the crossbred F1 females did not perform significantly better than outbred females for either total brood size or *w* (CRNOKRAK and BARRETT 2002). Block effects were not significant for any of the traits in *C. remanei*.

Over the course of serially inbreeding for 13 generations, 34 of 39 lines went extinct. To avoid the possibility of lines failing to mate successfully, we set up two plates of each inbred line in each generation. From the first generation of inbreeding, we found that 8.3% of single male-female crosses failed to mate in a 6 h period. Because we observed no difference between inbred and outbred crosses in the frequency of failure to mate in both the parental cross ($\chi_1^2 = 0.08$, $P = 0.78$) and the F1 × PB4641 cross ($\chi_1^2 = 0.05$, $P = 0.83$), we assume that failure is a random event; therefore, <1% of lines should fail to mate on both plates, although this is conservative since we allowed the inbred lines to mate indefinitely. Figure 4.4 shows the per-generation and cumulative rates of extinction of inbred lines. The per-generation extinction rate was consistently much greater than 1% (t_{12} = 4.00,

$P = 0.0009$), with a mean of 13.9% (SE = 3.2%), and there was a positive association between the per-generation rate of extinction and time, as indicated by a permutation test on a linear regression ($r^2 = 0.44$, $P = 0.015$). Since this significance value is nearly equivalent to that found for a parametric regression analysis ($t_{11} = 2.97$, slope = 0.020, $P = 0.013$), we used a parametric quadratic regression to test for purging. Purging of deleterious mutations over time would be expected to lead to a decreasing rate of extinction (CRNOKRAK and BARRETT 2002); however, a quadratic regression did not provide a better fit to the data ($t_{10} = 0.51$, $P = 0.62$), indicating that fitness decline does not taper off. It is worth noting that the extant lines after 13 generations were derived from all three initial sub-populations. Preliminary data from sequencing of exonic nuclear DNA indicates a large number of single nucleotide polymorphisms between the inbred lines, with some residual heterozygosity in at least one of the lines (results not shown). Thus, there appears to be no tendency to fix particular genotypes from a single sub-population.

4.4 Discussion

4.4.1 Outbreeding Depression in *C. elegans*

Outbreeding depression is defined as the reduction in fitness following hybridization between divergent populations (TEMPLETON 1986). When this occurs in first generation (F1) hybrids, as in our study of *C. elegans*, it could be caused by a disruption of beneficial interactions at three possible levels: between genes and the environment (local adaptation), within loci (underdominance), and between loci (coadapted gene complexes). The first two possibilities can probably be ruled out in our case. First, any effects of adaptation to local environmental conditions should be absent in the constant artificial laboratory environment, and there is no reason to suppose that outbred individuals should be any less adapted to this novel environment. Second, it is well-known that underdominance does not lead to the maintenance of variability, even with high selfing rates (KIMURA and OHTA 1971; ROCHELEAU and LESSARD 2000), unless there is frequency-dependent selection

under conditions similar to those involving local adaptation (WILSON and TURELLI 1986). Therefore, the observed pattern of outbreeding depression on fecundity can probably be attributed to a breakdown of coadapted gene complexes, revealing deleterious allele combinations affecting the fitness of the F1 hybrids or the viability of the F2 recombinant progeny (PHILLIPS and JOHNSON 1998). On the other hand, we did not find any inbreeding or outbreeding depression for longevity. This is consistent with earlier studies that found no effect of breeding class on life span (JOHNSON and WOOD 1982; JOHNSON and HUTCHINSON 1993), and little evidence of directional effects of new mutations on longevity (KEIGHTLEY and CABALLERO 1997; VASSILIEVA and LYNCH 1999; KEIGHTLEY *et al*. 2000; HALLIGAN *et al*. 2003).

We studied a limited number of crosses between strains from France; however, since *C. elegans* shows little or no geographic structuring of molecular diversity on a large scale (BARRIÈRE and FÉLIX 2005, 2007; HABER *et al.* 2005; CUTTER 2006), outbreeding depression is probably not limited to strains from France. Indeed, outbreeding depression appears to be a global phenomenon in *C. elegans*. In crosses between the canonical N2 strain and a Hawaiian strain, CB4856, the progeny of F1 hybrids exhibit a significant increase in embryonic and early larval lethality, with similar hybrid incompatibilities observed between other wild strains (SEIDEL *et al*. 2008; M. AILION, personal communication). In the related selfing species, *C. briggsae*, reproductive isolation was also found in crosses between strains AF16 and HK104, with approximately one third of F2 progeny exhibiting a delay in development and reduced intrinsic growth rates (S. BAIRD, unpublished results).

Previous studies failed to show any effect of inbreeding level on fitness-related traits in *C. elegans* (JOHNSON and WOOD 1982; JOHNSON and HUTCHINSON 1993; CHASNOV and CHOW 2002). However, as pointed out in the introduction to this paper, these all involved crosses where one of the strains was the long-maintained N2 strain, and the partner strains in the crosses had also been in culture for extended periods of time (HODGKIN and DONIACH 1997). Here we attempted to overcome this problem by using strains recently caught in the wild. Using similar laboratory techniques to previous studies, we found a significant reduction in the performance of outcrossed individuals. Obtaining estimates under standardized optimal conditions is an important first step, but it is uncertain how this effect translates to performance

under natural conditions. It is generally assumed that inbreeding depression is enhanced in harsher natural environments (see reviews by CRNOKRAK and ROFF 1999; KELLER and WALLER 2002; ARMBRUSTER and REED 2005). Whether this is also widely true of outbreeding depression remains to be seen, although in the selfing hermaphroditic snail, *Physa acuta*, outbreeding depression was observed only in the field but not in the lab (HENRY *et al*. 2003). Potential evidence of outbreeding depression and selection against hybrids in wild populations of *C. elegans* comes from recent regular monitoring of haplotype frequencies in a population in France, where the frequency of a recombinant haplotype decreased significantly over the span of a year (BARRIÈRE and FÉLIX 2007). Because the ecology of *C. elegans* is poorly understood, data from more natural conditions are needed.

4.4.2 Implications for *C. elegans* Diversity

It has long been recognized that selfing can promote the maintenance of linkage disequilibrium among loci whose alleles have epistatic effects on fitness ("coadaptation") (e.g. STEBBINS 1957; ALLARD 1975), although selfing can also enhance randomly-generated linkage disequilibrium between neutral variants, because it reduces the effective rate of recombination (CHARLESWORTH 2003). In *C. elegans*, extremely low levels of diversity are found, with extensive linkage disequilibrium across the genome, both within and between chromosomes (KOCH *et al*. 2000; BARRIÈRE and FÉLIX 2005, 2007; HABER *et al*. 2005; CUTTER 2006). This pattern is thought to be determined largely by self-fertilization and population subdivision, combined with rare but regular migration and outcrossing.

Our results suggest that outbreeding depression could be an additional important factor in shaping the genetic structure of *C. elegans*. Selection against hybrids would result in a reduction in the effective migration and outcrossing rates (BARTON and BENGTSSON 1986), thereby maintaining higher levels of linkage disequilibrium than in its absence. Outbreeding depression may also help to explain the large discrepancies between estimates of outcrossing rates based on heterozygote frequencies at microsatellite loci (BARRIÈRE and FÉLIX 2005, 2007; SIVASUNDAR and HEY 2005), and measures from linkage disequilibrium, which are typically one to

three orders of magnitude lower (BARRIÈRE and FÉLIX 2005, 2007; CUTTER 2006). If selection maintaining coadapted gene complexes caused hybrids between strains to have reduced contribution to the next generation, then the effective outcrossing rate will be greatly reduced.

4.4.3 Frequency of Males

Because males are non-essential for reproduction in *C. elegans*, their maintenance requires explanation. It is possible that androdioecy could simply be a byproduct of the sex-determination mechanism, which allows male production via non-disjunction of X-chromosomes in hermaphrodite meiosis (hermaphrodites and females are XX, males are XO). CHASNOV and CHOW (2002) argued that males are non-adaptive but persist because mating is sufficiently frequent to preclude degeneration of male-specific genes by deleterious mutation. Other arguments favour a selective advantage of outcrossing to explain the persistence of males, and it is conceivable that males and outcrossing are advantageous under different ecological settings (SCHULENBURG and MÜLLER 2004), or particular demographic scenarios involving metapopulation dynamics (PANNELL 2002; WEEKS *et al.* 2006a). Our results, however, suggest that males and outcrossing will be selected against in natural populations, since the existence of outbreeding depression must reduce the frequency of males below that expected in its absence. This was verified by using Equation (3*a*) of CUTTER *et al.* (2003). For example, with the male reproductive efficiency calculated for the N2 strain (CUTTER *et al.* 2003), which has a fairly typical male mating ability for the species (TEOTÓNIO *et al.* 2006), the observed level of outbreeding depression of 6–9% would reduce the expected equilibrium frequency of males by about a third for any rate of X chromosome non-disjunction.

4.4.4 Inbreeding Depression in *C. remanei*

Consistent with expectations (e.g. LANDE and SCHEMSKE 1985; CHARLESWORTH and CHARLESWORTH 1998), the outcrossing species *C. remanei* exhibited much greater levels of inbreeding depression than the selfing *C. elegans*.

This appears to reflect the effects of inbreeding on various components of fitness. *C. remanei* female reproduction is limited by the number of sperm transferred by their male partners, which is a product of sperm production and mating efficiency. We attempted to standardize the male contribution by allowing mating for a controlled period of time. We also ensured that mating had indeed occurred by discarding any crosses for which no progeny were observed, although this might underestimate inbreeding depression by ignoring crosses that were unsuccessful due to genetic incompatibilities rather than a failure to copulate. Also, by mating the F1 female progeny to PB4641 males, sperm-limitation and male mating efficiency should be constant and independent of inbreeding level. This allowed us to compare directly the performance of inbred and outbred females. In these crosses, inbreeding depression will be solely due to females, whereas without the control PB4641 males, inbreeding depression can also be affected by any male component and by larval inviability of offspring. Comparisons of parental crosses with F1 × PB4641 crosses indicate that inbreeding depression in *C. remanei* is affected by both male and female performance, in addition to larval inviability (see Figure 4.3).

4.4.5 Evidence of Maternal-Effect Inbreeding Depression

The result that crossbred F1 females mated to PB4641 males had lower *w* than outbred females, despite having nearly the same brood sizes, suggests that outcrossing does not completely eliminate the effect of prior inbreeding on fitness in *C. remanei*. This indicates a maternal effect on inbreeding depression. A few other studies have found that maternal inbreeding can influence outcrossed progeny fitness, although the magnitude of these effects is generally less than the direct genetic effects of inbreeding depression (e.g. HAUSER and LOESCHCKE 1995; LYONS 1996; VOGLER *et al*. 1999; HAYES *et al*. 2005). This highlights the need to separate the sources of inbreeding depression, in order to accurately estimate the magnitude of the direct genetic effects. Our initial F1 × PB4641 crosses should not have been influenced by maternal inbreeding depression, since the parents were outbred, so these should provide a good estimate of zygotic inbreeding depression.

The significantly lower *w* suggests that it is the timing of egg production and fertilization rather than the absolute brood size that is altered by maternal-effect inbreeding depression. In *C. elegans*, a number of maternal-effect genes have been identified that influence developmental and behavioural timing, known as Clk genes (LAKOWSKI and HEKIMI 1996). For example, *clk-1* mutants have a two- to six-fold reduction in their egg production rate, but display maternal rescue, indicative of maternal effects (WONG *et al*. 1995). It is possible that some of the inbred *C. remanei* lines in our experiment became homozygous for genes with similar effects, thereby causing maternal-effect inbreeding depression on physiological traits related to the timing of the reproductive schedule in the crossbred F1 progeny.

4.4.6 Evolution of Androdioecy

Recent evidence suggests that the genetics of switching between different reproductive modes is relatively simple in *Caenorhabditis* species, requiring a very limited number of mutations (NAYAK *et al*. 2005; BRAENDLE and FÉLIX 2006; HILL *et al*. 2006). This demonstrates the potential for frequent evolutionary transitions in mating systems in *Caenorhabditis*. Selfing could have many advantages in *Caenorhabditis*, most notably reproductive assurance in species colonizing ephemeral habitats with fluctuating population dynamics and episodic low densities, although the natural history of the genus is poorly understood and it is unclear exactly how the habitats differ between the gonochoristic and androdioecious species (KIONTKE and SUDHAUS 2006; WEEKS *et al*. 2006a). The potential benefits of selfing, however, must be weighed against the cost of overcoming the strong inbreeding depression found in *C. remanei*, which presumably reflects conditions experienced by a lineage upon first shifting to hermaphroditism.

Over the course of 13 generations of full-sibling mating in *C. remanei*, 87% of our inbred lines went extinct and the surviving lines were extremely sick, with very low fecundity, and exhibited both direct and maternal-effect inbreeding depression (see Figures 4.3 and 4.4). Among the extant lines, there was little evidence of any rebound in fitness over time that would have indicated purging of deleterious alleles (CRNOKRAK and BARRETT 2002), which suggests that the

inbreeding depression is largely due to the cumulative effects of many deleterious alleles of small effect rather than a few segregating deleterious alleles of large effect (WILLIS 1999). Indeed, most inbred strains of *C. remanei* die out in the laboratory, and those that survive have much lower fecundity and longer generation times than non-inbred strains, and appear to be less fit than recently established strains (S. BAIRD & E. DOLGIN, unpublished observations). Furthermore, newly arising hermaphroditic lineages would probably have suboptimal sperm numbers and brood sizes, thereby decreasing the rate of population expansion (CUTTER 2004), lowering the effective population size, and increasing the likelihood of extinction.

In the long run, selfing lineages that survive are likely to be those that have purged much of their genetic loads; a higher degree of population subdivision in selfing lineages may facilitate such purging (WALLER 1993; THEODOROU and COUVET 2002; WHITLOCK 2002; GLÉMIN *et al*. 2003). This makes it hard for transitions back to outcrossing to occur (CHARLESWORTH and CHARLESWORTH 1998). But are transitions to hermaphroditism in rhabditid nematodes evolutionary dead-ends? Phylogenetic evidence suggests that hermaphroditism has evolved frequently in rhabditids, but these lineages rarely survive in the long run (KIONTKE and FITCH 2005). This is consistent with the pattern found in several plant genera (SCHOEN *et al*. 1997; GOODWILLIE 1999; TRUYENS *et a*l. 2005; but see TAKEBAYASHI and MORRELL 2001). On the other hand, in the clam shrimp *Eulimnadia*, androdioecy appears to be the ancestral mode of reproduction in the genus, and this mating system has persisted for between 24 and 180 million years (WEEKS *et al*. 2006b). However, the rate of selfing seems to be lower in *Eulimnadia*, and the level of inbreeding depression and maximal frequencies of males higher, than in *C. elegans* (WEEKS *et al*. 1999; WEEKS 2004). It is unknown how long *C. elegans* and *C. briggsae* have been selfing. Although the selfing species are separated by millions of years from their closest known relatives, the ecology and biodiversity of *Caenorhabditis* species is poorly understood and extensive sampling may yet identify new closely related outcrossing and self-fertile species, and narrow down the possible time since hermaphroditism evolved. Here, we have shown that natural isolates of *C. elegans* suffer from outbreeding rather than inbreeding depression. How the species arrived at this state, the time-scale of this process, and whether the species will persist are intriguing problems.

5 Genetic and Phenotypic Diversity in African Worms

The work presented in this chapter was published in**:**

- DOLGIN, E. S., M.-A. FÉLIX, and A. D. CUTTER, 2008 Hakuna Nematoda: genetic and phenotypic diversity in African isolates of *Caenorhabditis elegans* and *C. briggsae*. *Heredity*, advance online publication 12 December 2007, doi: 10.1038/sj.hdy.6801079.

Contributing authors:

- I collected the African strains, carried out the molecular assays and hybrid crosses, performed the analysis, and wrote the manuscript.
- M.-A. FÉLIX performed the vulval cell lineage analysis, took the Nomarski micrograph images, and collected some non-African strains.
- A. CUTTER advised on the projest, helped with the DNA isolation, designed the PCR primers, and assisted in writing the manuscript.

5.1 Introduction

As a model organism, the nematode *Caenorhabditis elegans* has provided an outstanding system for elucidating fundamental problems in biology. However, the attitude regarding the natural history and ecology of *Caenorhabditis* has long been "hakuna matata", to use the Swahili phrase meaning "no worries". To achieve a robust understanding of the evolutionary context in which *C. elegans*' developmental and genetic patterns emerged, it is necessary to characterize representative natural populations on a worldwide scale and to conduct comparative analyses of *Caenorhabditis* species. Specifically, the quantification of genetic and phenotypic

variation in species with shared common ancestry permits inference about the evolutionary forces and population processes that shape the life history and development of these organisms.

Along with *C. elegans*, *C. briggsae* forms part of a monophyletic clade of five species in laboratory culture known as the *Elegans* group of the genus *Caenorhabditis* (KIONTKE and FITCH 2005). Like *C. elegans*, *C. briggsae* is androdioecious (self-fertile hermaphrodites and facultative males), but both species evolved independently from gonochoristic (male–female) ancestors (KIONTKE *et al.* 2004). Despite analogous breeding systems and similar morphology (NIGON and DOUGHERTY 1949), their genome sequences have diverged drastically (COGHLAN and WOLFE 2002; STEIN *et al.* 2003), although they do show a high degree of chromosomal synteny (HILLIER *et al.* 2007). The development of *C. briggsae* as a laboratory system is still in its infancy, but with a growing number of genetic resources becoming available, including a sequenced genome and recombination maps, this species is emerging as a useful companion species of *C. elegans* (BAIRD and CHAMBERLIN 2006).

C. elegans and *C. briggsae* are cosmopolitan species with overlapping geographic distributions, with at least one of the two species found on each continent where intensive field collections have been carried out. Both species show similarly low levels of genetic variation, especially when compared to the related gonochoristic species, *C. remanei* (GRAUSTEIN *et al.* 2002; JOVELIN *et al.* 2003; HAAG and ACKERMAN 2005; CUTTER 2006; CUTTER *et al.* 2006a,b) . The predominantly self-fertilizing breeding systems shared by both *C. elegans* and *C. briggsae* has probably driven this decrease in diversity as a consequence of reduced effective population size, increased homozygosity, and greater population subdivision (CHARLESWORTH 2003). Despite their similarities, however, *C. elegans* and *C. briggsae* probably occupy different ecological niches. For example, one difference relevant to ecology is that *C. briggsae* tolerates and proliferates at higher temperatures than *C. elegans* (FODOR *et al.* 1983). *C. elegans* is also strongly induced to form the dormant dauer larval stage at high temperatures, whereas *C. briggsae* appears to lack this response (INOUE *et al.* 2007).

Another difference of potential ecological significance is the relative importance of geography in structuring genetic variation. Phylogenetic reconstructions of molecular polymorphisms among available global samples show no strong signature of geographic structure in *C. elegans* (DENVER *et al.* 2003; BARRIÈRE and FÉLIX 2005; HABER *et al.* 2005; CUTTER 2006), but a clear division is observed in *C. briggsae* between strains from temperate latitudes and strains found near the Tropic of Cancer (GRAUSTEIN *et al.* 2002; CUTTER *et al.* 2006b). A major problem with our current understanding of natural variation and biogeography, however, is that it is not clear whether the focal samples adequately reflect global diversity, because sampling efforts have not targeted most regions outside North America and Europe. Although the genus was first described from specimens isolated in North Africa (MAUPAS 1900), these strains were never cultured, and, to date, no *C. elegans* or *C. briggsae* strains have been analyzed from Africa or South America. To address this issue, one of us (E.S.D.) undertook intensive field collections in two countries in sub-Saharan Africa: Kenya and South Africa.

In this study, we test the hypotheses concerning geographic structuring of genotypes in *C. elegans* and *C. briggsae* with new samples from Africa. We also examine vulval cell lineages for phenotypic polymorphisms and conduct intra-specific crosses in *C. briggsae* to test for genetic incompatibilities. Our results confirm the lack of global geographic structure in the sequences of *C. elegans*, but provide further evidence for latitudinal clades in *C. briggsae*. We also find unique patterns of vulval cell division in South African strains of *C. briggsae* that have not been observed previously in this species, while crosses between isolates from disparate localities fail to demonstrate hybrid incompatibilities (DOLGIN *et al.* 2007), and instead show evidence of inbreeding depression. We discuss how these new isolates add to our understanding of global variation and how evolutionary processes might act differently on the two species after their independent transitions to androdioecy.

5.2 Materials and Methods

5.2.1 Nematode Populations

We isolated 38 new strains of *C. elegans* and 25 new strains of *C. briggsae* from single wild hermaphrodites in Kenya and South Africa in March–April 2006. Despite extensive sampling from 36 sites, *Caenorhabditis* nematodes were only found in four locations: (1) compost from a private garden (Johannesburg, South Africa); (2) compost from a plant nursery (Ceres, South Africa) (3) compost from a mushroom farm (Limuru, Kenya); and, (4) leaf litter in a public park (Nairobi, Kenya). Strain name designations for *C. elegans* are ED3040–ED3077, and for *C. briggsae* are ED3078–ED3102. The sampling sites and strain designations are described in Table 5.1. Various other nematode species were also found in these sites, including many strains of *Oscheius tipulae*, which have also been characterized (BAÏLLE *et al*. 2008).

The sampling protocol was done using the "Isolation on an Agar Plate" technique described in BARRIÈRE and FÉLIX (2006). Briefly, small samples of compost or leaf litter were placed on 9-cm NGM-lite agar plates spotted with *Escherichia coli* OP50. These plates were then monitored for nematodes crawling out into the bacterial lawn for 24–48 hours, and individual worms were isolated onto separate 6-cm plates to establish independent iso-hermaphrodite lineages. Self-fertile individuals were examined for morphological features under compound microscopy and species identity was made through mating tests and by sequencing a portion of the small sub-unit ribosomal RNA gene (FLOYD *et al.* 2002).

For DNA sequencing, we randomly chose a subset of 16 strains of each species from the African collection as representative of these sampling sites (Table 5.1). The sequences from these strains were then compared to isolates from a larger worldwide data set of 118 *C. elegans* strains (CUTTER 2006) and 63 *C. briggsae* strains (CUTTER *et al.* 2006b). To gain a more complete understanding of the molecular diversity in the two species, we also used an additional nine strains that were recently isolated from geographically unique locations and had not previously been assessed for nucleotide polymorphism: a *C. elegans* strain from Madagascar

Table 5.1: New *C. elegans* and *C. briggsae* strains isolated in Kenya and South Africa, and a list of the strains used for DNA sequencing. *C. elegans* haplotype designations are as in Figure 5.2A.

Location (nearest city)	Latitude (S)	Longitude (E)	Altitude (m)	*C. elegans*			*C. briggsae*		
				#	Strain designations	Strains used in sequencing (and haplotypes)	#	Strain designations	Strains used in sequencing
Johannesburg, South Africa	26°10'	28°01'	1669	1	ED3040	ED3040 (α)	12	ED3078–ED3089	ED3078 ED3079 ED3080 ED3082 ED3083 ED3084 ED3087 ED3089
Ceres, South Africa	33°22'	19°19'	478	12	ED3041–ED3052	ED3042 (β) ED3046 (γ) ED3048 (β) ED3049 (γ) ED3051 (β) ED3052 (δ)	0		
Limuru, Kenya	1°05'	36°39'	2356	24	ED3053–ED3076	ED3054 (ε) ED3057 (ε) ED3063 (ε) ED3066 (ε) ED3072 (ε) ED3073 (ε) ED3075 (ε) ED3076 (ε)	0		
Nairobi, Kenya	1°19'	36°48'	1708	1	ED3077	ED3077 (β)	13	ED3090–ED3102	ED3091 ED3092 ED3095 ED3096 ED3098 ED3099 ED3100 ED3101

(LKC34; isolated by V. STOWELL and donated by L. CARTA), a *C. elegans* strain from Ribeiro Frio, Madeira (JU258; isolated by M.-A. FÉLIX), a *C. elegans* strain from Kakegawa, Japan (JU1088; isolated by M.-A. FÉLIX), two *C. elegans* strains from Concepcion, Chile (JU1171 and JU1172; isolated by M.-A. FÉLIX), a *C. briggsae* strain from Kruger National Park, South Africa (DF5100; isolated by W. SUDHAUS and donated by K. KIONTKE), a *C. briggsae* strain from Kakegawa, Japan (JU1085; isolated by M.-A. FÉLIX), and two *C. briggsae* strains from Salt Lake City, Utah, USA (EG4181 and EG4207; isolated by M. AILION and obtained from the Caenorhabditis Genetics Center). The geographic origin of each *C. elegans* and *C. briggsae* strain included in this analysis is shown in Figure 5.1.

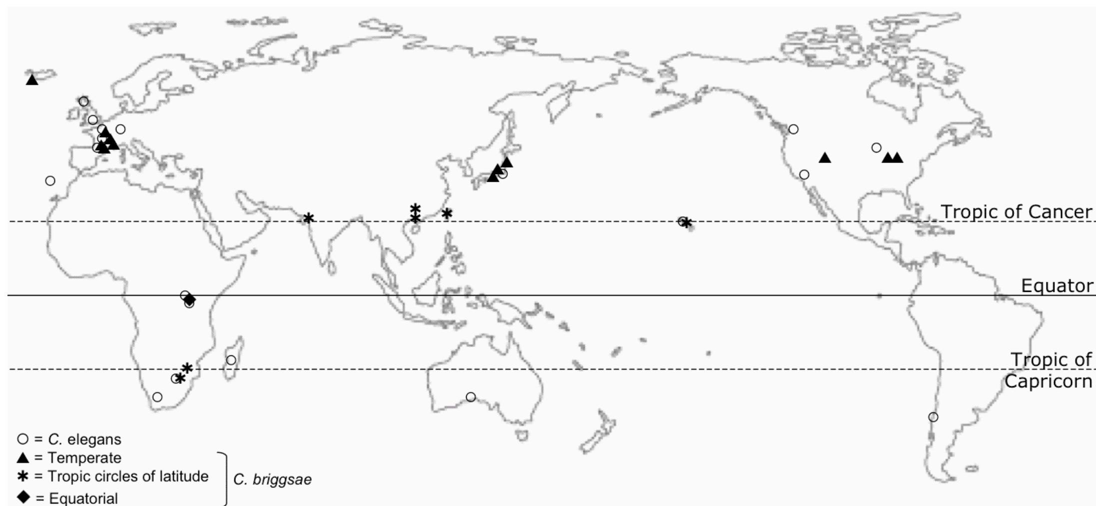


Figure 5.1: Geographic distribution of strains used for molecular analysis. Open circles indicate *C. elegans* strains and closed symbols indicate *C. briggsae* strains from different latitudinal clades: triangles (temperate latitude samples), asterisks (Tropic circles of latitude samples), and diamonds (equatorial samples). Note that the exact sampling location of *C. elegans* strain LKC34 is uncertain although it is known to originate from Madagascar.

5.2.2 Molecular Methods

DNA from single individuals was obtained using a NaOH digestion protocol (FLOYD *et al.* 2002). We sequenced the gene fragments that had previously been analyzed for a larger global collection of strains (CUTTER 2006; CUTTER *et al.* 2006b). These comprised six genes from *C. elegans* chromosomes II and X, and the putative orthologs of five of these genes plus one locus nearby the sixth gene in *C. briggsae*, with primers designed to span a long (>500 bp) intron. Names, lengths, positions and primers of the sequenced loci can be found in CUTTER (2006) and CUTTER *et al.* (2006b). Both strands were sequenced on an ABI Prism 3730 automated sequencer at the School of Biological Sciences Sequencing Service, University of Edinburgh.

5.2.3 Sequence Analysis

Sequence alignment and manual editing to remove the primers were performed with Sequencher 4.6 and BioEdit 7.0.5. Calculations of diversity (from pairwise differences, π, and from the number of segregating sites, θ), Tajima's *D* tests of neutrality (TAJIMA 1989), and tests of population differentiation using average values of K_{st}^{*} among loci (HUDSON *et al.* 1992) and Hudson's (HUDSON 2000) nearest-neighbor statistic, S_{nn}, from concatenated sequences were made using DnaSP 4.10.9 (ROZAS *et al.* 2003). We present diversity data from silent sites (synonymous and intronic positions); sites corresponding to indels or incomplete data were excluded from the analyses. Neighbor networks and neighbor-joining trees were constructed with concatenated sequences using SplitsTree 4.6 (HUSON and BRYANT 2006).

5.2.4 Vulval Cell Lineage

The frequency of division of the vulval precursor cells, P3.p and P4.p, was measured in a subset of the new African isolates—5 *C. elegans* strains (ED3040, ED3046, ED3052, ED3054, and ED3077) and 4 *C. briggsae* strains (ED3082, ED3087, ED3092, and ED3101)—in addition to *C elegans* strains JU258 and LKC34. The results were compared to a global collection of 13 *C. elegans* strains and 6 *C. briggsae* strains that had previously been scored for these polymorphic phenotypes (DELATTRE and FÉLIX 2001b). For some of these previously analyzed strains, more worms were also scored to obtain larger sample sizes. We also screened eight strains from other species in the *Elegans* group of the *Caenorhabditis* genus for comparison—four strains of *C. remanei* (JU724, JU825, PB4641 and SB146), two strains of *C. brenneri* (CB5161 and PB2801), and two strains of *C.* sp. 5 (JU727 and SB378). The fate of the Pn.p cells was assessed at the L4 stage using Nomarski miscroscopy under standard conditions. We tested for heterogeneity between intraspecific strains by using replicated goodness-of-fit tests on the number of individuals in which the Pn.p cell divided or not, and performed post-hoc pairwise tests after applying the Dunn-Šidák correction (SOKAL and ROHLF 1995, p. 239).

5.2.5 Hybrid Crosses

To test for possible incompatibilities between strains of *C. briggsae*, a series of crosses were set up using strains from South Africa (ED3083) and Kenya (ED3101), as well as three additional *C. briggsae* strains: AF16 from India, HK104 from Japan, and ED3034 from Taiwan. Pairwise crosses to the two African strains created seven crosses: ED3083 × ED3034, ED3083 × AF16, ED3083 × HK104, ED3083 × ED3101, ED3101 × ED3034, ED3101 × AF16, and ED3101 × HK104. We also performed an eighth cross between AF16 and HK104 since this inter-strain cross has been studied previously (BAIRD *et al.* 2005). Hybrid crosses were only performed in *C. briggsae* in this study because only African samples of this species showed markedly different molecular and developmental patterns (see results), and hybrid crosses with representative populations have been performed elsewhere for *C. elegans* (DOLGIN *et al.* 2007).

In each cross, we generated mixed-mating populations of males (M) and hermaphrodites (H) for each strain. We then established pure strain F1s from each strain (e.g. ED3101 × ED3101, and AF16 × AF16), and the two reciprocal F1 hybrids (e.g. ED3101–M × AF16–H, and AF16–M × ED3101–H). We then measured the number of surviving late larval F2s (to give a measure of F1 brood size), the percentage of F2 embryonic lethality, and the developmental timing to reach the L4 larval stage in the F2s. Inter-strain crosses of *C. briggsae* were performed using the methods that DOLGIN *et al.* (2007) used for *C. elegans*, with a few notable exceptions. Following synchronization of the pure-strain parental worms using alkaline hypochlorite, some L4 hermaphrodites were transferred to a new plate and the remaining worms were left with the males in a mixed-sex population. The isolated hermaphrodites were permitted to self for ~3.5 d following first egg-lay to exhaust them of self-sperm. Next generation males were then taken from the mixed-sex population and were mated to these sperm-depleted hermaphrodites for a 24 h interval. The mated hermaphrodites were subsequently transferred to new plates to lay eggs for 5 h. From these F1 progeny, many L4 hermaphrodites of each pure-strain or hybrid genotype (mean number = 19.2 ± 1.1 SE) were set up individually to

measure surviving larval brood size and embryonic lethality. These worms were transferred every 24 h to a new plate; 36 h after transfer, unhatched eggs and surviving larvae were counted. Subsets of the remaining F1 hermaphrodites were transferred en masse to new plates at the L4 stage. Several hours after eggs were observed on these plates, the worms were transferred twice again to new plates to lay eggs for two 1 h intervals. These plates were then monitored for when the F2 progeny molted from the L4 stage to adulthood, as a measure of developmental timing. The mean brood sizes and levels of embryonic lethality within the same hybrid cross were compared using *t*-tests. Differences in developmental timing were evaluated using goodness-of-fit tests on the number of "normal" or "late" molting larvae, where late was defined as more than 5h after the peak molting time.

5.3 Results

5.3.1 Molecular Polymorphism

The African sample of *C. elegans* shows similar low levels of within-locality variation as reported previously for other regions of the world. Overall silent-site nucleotide polymorphism in the African localities was equivalent to European samples, with Africa-wide π_{si} and θ_{si} both estimated to be ~0.2% (Table 5.2). Considering diversity levels of South African or Kenyan isolates independently also resulted in quantitatively similar diversity levels to that found in individual European localities. Despite low levels of variation, there is no evidence of non-neutral demographic or selective processes in the African samples, either as a whole or by considering each country individually, as there was a range of positive and negative values of Tajima's *D* across the different loci, with no values significantly different from zero (results not shown).

The 17 African samples of *C. elegans* (including LKC34 from Madagascar) contained six different haplotypes. Although these haplotypes were endemic to African samples, most polymorphic sites are also found elsewhere in the world (Figure 5.2A). All the African strains were generally similar to other strains and

Table 5.2: Summary of nucleotide polymorphisms for African *C. elegans* and *C. briggsae* isolates. Diversity estimates were calculated for each country in both species, and for all African samples for *C. elegans* but not for *C. briggsae* due to the distinct geographical structuring between countries in this species. For *C. elegans*, there were 17 strains total: 7 from South Africa, 9 from Kenya, and 1 from Madagascar (LKC34). For *C. briggsae*, there were 17 strains total: 9 from South Africa (including DF5100), and 8 from Kenya.

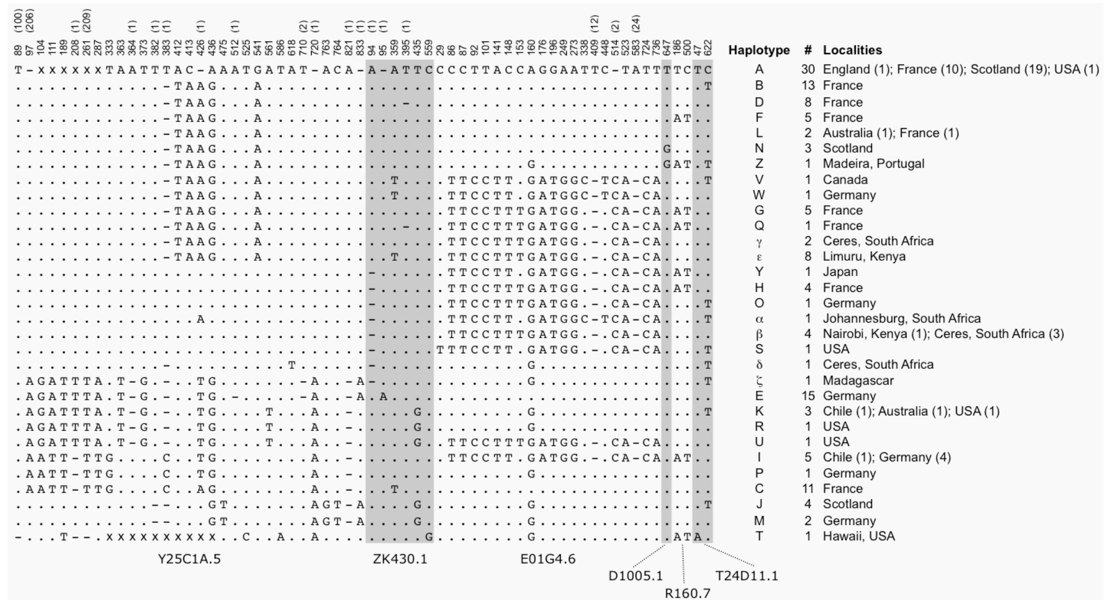
		African population samples							***South Africa***		***Kenya***	
	Locus (chromosome)	**Silent sites**	***S***	***h***	H_d	**Indels**	π_{si}	θ_{si}	π_{si}	θ_{si}	π_{si}	θ_{si}
C. elegans	Y25C1A.5 (II)	658.8	7	4	0.596	4	0.0036	0.0031	0.0033	0.0031	0.0013	0.0022
	ZK430.1 (II)	542.8	1	2	0.529	1	0.0010	0.0005	0	0	0.0004	0.0007
	E01G4.6 (II)	521.4	16	3	0.324	2	0.0066	0.0091	0.0092	0.0125	0	0
	D1005.1 (X)	480.0	0	1	0	0	0	0	0	0	0	0
	R160.7 (X)	605.2	0	1	0	0	0	0	0	0	0	0
	T24D11.1 (X)	651.0	1	2	0.309	0	0.0005	0.0005	0.0007	0.0006	0	0
	Concatenated	3459.2	25	6	0.743	7	0.0019	0.0021	0.0021	0.0026	0.0003	0.0005
	Average	576.5	4.2	2.2	0.293	1.2	0.0019	0.0022	0.0022	0.0027	0.0003	0.0005
C. briggsae	p09	484.7							0	0	0	0
	p10	581.3							0.0004	0.0006	0.0004	0.0007
	p11	571.3							0	0	0	0
	p12	623.7							0.0007	0.0012	0	0
	p13	593.5							0	0	0	0
	p14	599.5							0	0	0	0
	Concatenated	3452.0							0.0002	0.0003	0.0001	0.0001
	Average	575.7							0.0002	0.0003	0.0001	0.0001

S, segregating sites; *h*, haplotypes, H_d, haplotype diversity; π_{si} , silent-site diversity from pairwise differences; θ_{si}, silent-site diversity from the number of segregating sites.

showed no strong patterns of geographic structure (Figures 5.3); however, we did find some evidence of population differentiation between European and African samples (K_{st}^{*}= 0.075; S_{nn} = 0.97, $P < 0.0001$) and between Kenyan and South African samples (K_{st}^{*}= 0.215; S_{nn} = 0.85, $P = 0.003$). Nonetheless, after including the African samples, we did not find evidence of any additional recombination events to those found previously, as indicated by the four-gamete test (HUDSON and KAPLAN 1985). We also sequenced loci for the first known *C. elegans* isolates from Asia (JU1088 from Japan) and South America (JU1171 and JU1172 from Chile), plus the Madeiran strain JU258, since it was previously reported that this strain was highly differentiated from other strains (HABER *et al.* 2005; STEWART *et al.* 2005). Despite the unique geographical origin of these additional strains, we failed to find high divergence of any of them from global samples for these six loci (see haplotype Y for JU1088, haplotype I for JU1171, haplotype K for JU1172, and haplotype Z for JU258, in Figures 5.2A and 5.3). Furthermore, JU258 was not strongly allied to the Hawaiian strain, CB4856 (haplotype T), relative to the N2 strain (haplotype A), as previously suggested (HABER *et al.* 2005; STEWART *et al.* 2005; MAYDAN *et al.* 2007).

The African *C. briggsae* samples show a markedly different pattern from that found for *C. elegans*. We observed very little within-locality diversity across the six loci tested, with South African and Kenyan π_{si} and θ_{si} both estimated to be extremely low (Table 5.2). This results from the fact that 7 out of 8 Kenyan strains were identical, with only a single SNP distinguishing ED3092, and of the 9 South African strains, the 8 strains from Johannesburg were identical to each other, and DF5100, from Kruger National Park, had only three SNP differences from the Johannesburg strains. Nonetheless, the African *C. briggsae* haplotypes were very informative about geographic structure (Figure 5.2B). The South African haplotypes were remarkably similar to those of other strains found in the northern hemisphere near the Tropic of Cancer, and the Kenyan haplotypes were unlike those of any other strains previously described with many newly identified polymorphisms not shared with other populations (Figure 5.4). For example, the p11 locus had previously been shown to be monomorphic (CUTTER *et al.* 2006b), but we found 7 SNPs and 3 indels in the Kenyan samples (Figure 5.2B). However, since most of these polymorphisms were

A



B

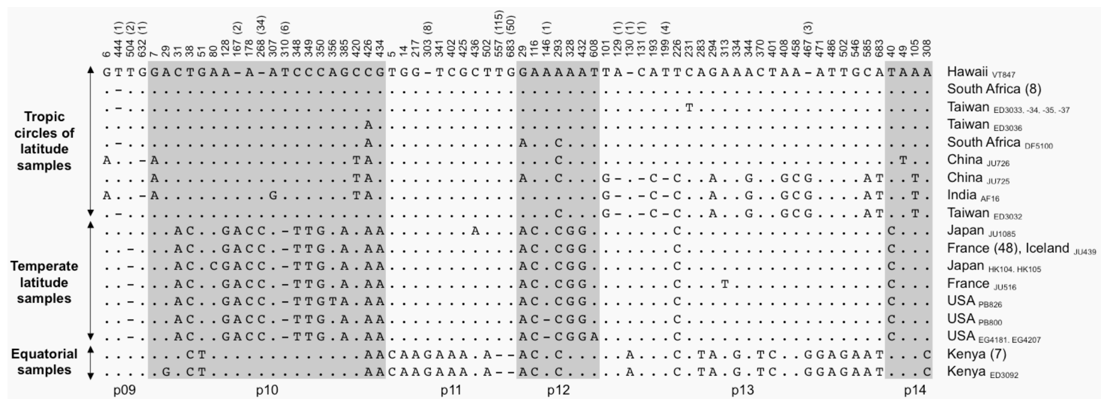


Figure 5.2: Summary of multilocus haplotypes for *C. elegans* (A) and *C. briggsae* (B). Loci are labeled along the bottom, with the positions of polymorphisms relative to the start of each locus indicated across the top (indel lengths in parentheses). For *C. elegans*, the haplotype designations follow those of CUTTER (2006), with JU1088 denoted by haplotype Y, JU258 denoted by haplotype Z and the African haplotypes indicated by Greek letters. For a list of which strains correspond to each African haplotype, see Table 5.1. Haplotype X of CUTTER (2006) was found to be equivalent to A, as a consequence of an alignment error at position 426 in Y25C1A.5. We also amend several presentation errors from Figure 1 of CUTTER (2006).

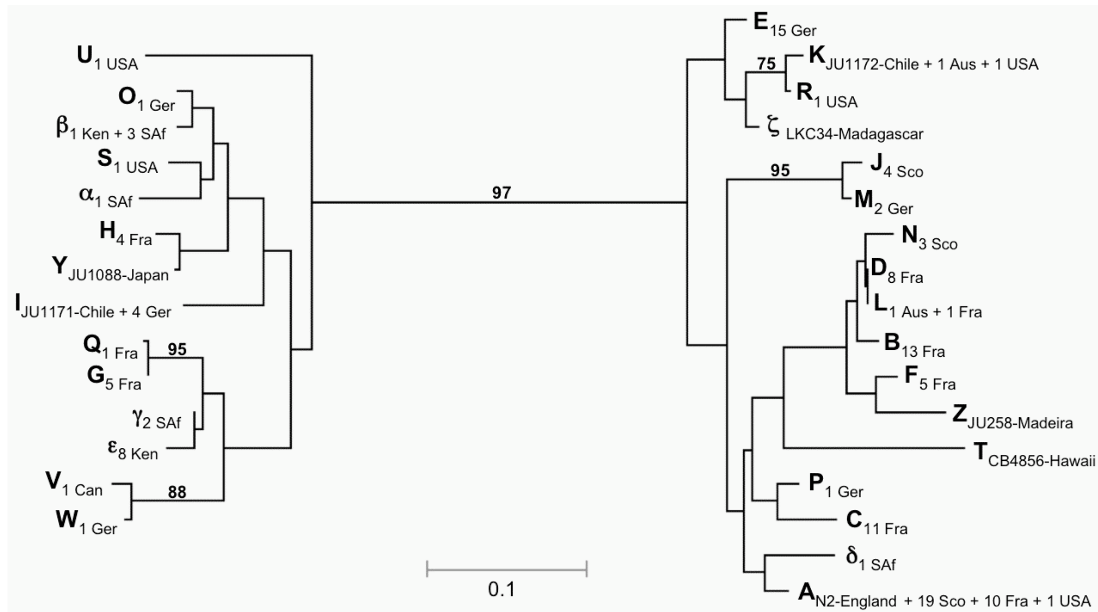
unique to this locality, we found no further evidence of historical recombination after including the African samples, as indicated by the four-gamete test (HUDSON and KAPLAN 1985).

The "tropical" strains previously described originated from China, India and Hawaii—all situated within 3° of the Tropic of Cancer. The South African strains described here were isolated in Johannesburg, less than 3° south of the Tropic of Capricorn, and Kruger National Park, which straddles the Tropic of Capricorn. Considering that these samples were all quite different from the Kenyan samples, which is also technically a tropical location, we propose that what was previously called "tropical" strains (CUTTER *et al.* 2006b) should be considered as "Tropic circles of latitude" strains, and suggest that molecular diversity in *C. briggsae* strains may be partitioned into three major latitudinal clades: temperate samples, Tropic circles of latitude samples, and equatorial samples (Figures 5.2B and 5.4). An alternative model of geographic structure is that a variety of deeply divergent clades of *C. briggsae* populate the globe in a manner that only mimics a latitudinal distribution, as a consequence of available sampling. We also sequenced three additional recently isolated strains: a strain from Kakegawa, Shizuoka Prefecture, Japan (JU1085), and two from Salt Lake City, Utah, USA (EG4181 and EG4207), and observed that these strains were quite similar to other strains from temperate latitudes, consistent with the prediction of a temperate latitudinal clade (see Figures 5.2B and 5.4). More extensive global sampling is required to confirm the generality of this latitudinal geographic structure to genetic diversity in *C. briggsae*.

5.3.2 Vulval Cell Lineage

In *C. elegans*, the significant variation we observed among strains in the division frequency of the P3.p vulval precursor cell (χ^2_{19} = 447.5, P < 0.0001) indicates that considerable heritable phenotypic diversity exists in the African sample, despite a lack of substantial genetic variation at the loci surveyed here (Figure 5.5). For example, ED3046 (South Africa) and ED3054 (Kenya) exhibited a 38% difference in P3.p division frequency, yet differed by only one SNP across the ~4kb molecular dataset (see Figure 5.2A). Unlike the variation known for P3.p cell

A



B

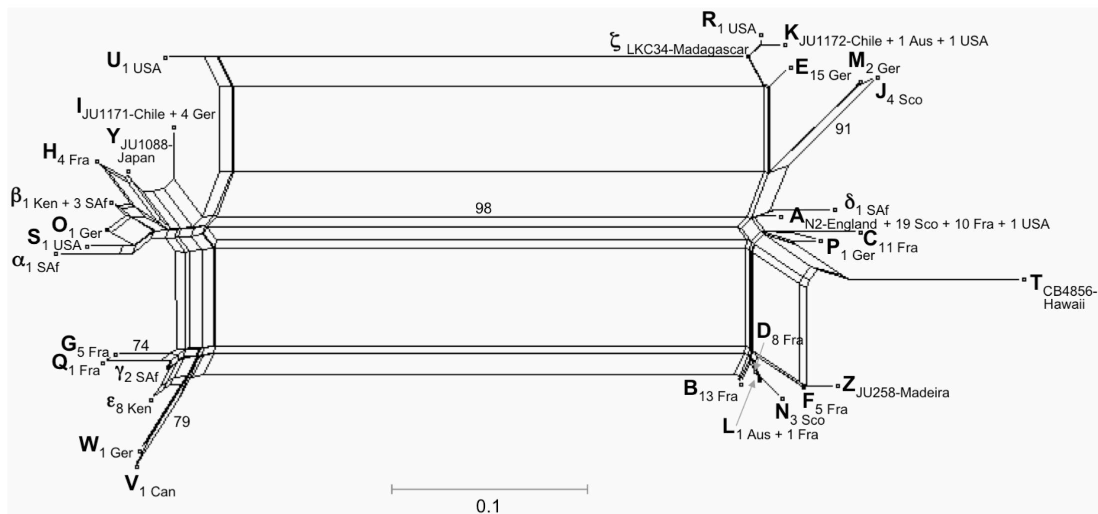
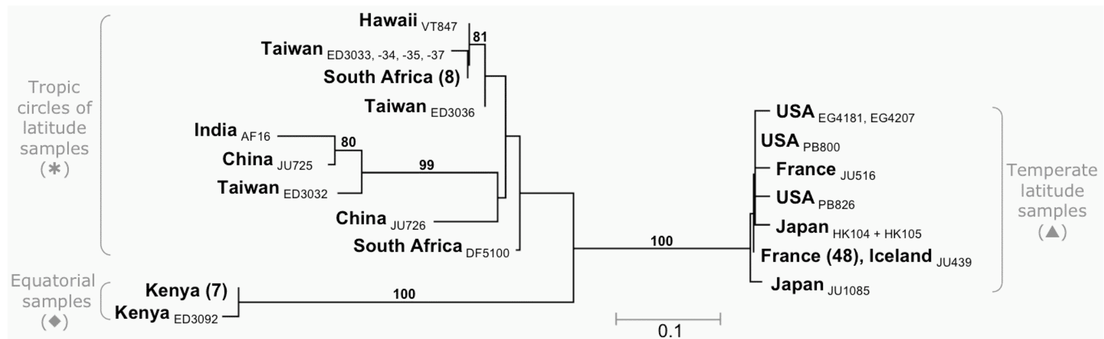


Figure 5.3: Unrooted *p*-distance neighbor-joining tree (A) and neighbor-net (B) for *C. elegans* based on concatenated multilocus haplotypes. Haplotype designations are as in Figure 5.2A; African haplotypes are designated by Greek letters; subscripts indicate the number of strains per haplotype and the country of origin (Aus, Australia; Can, Canada; Fra, France; Ger, Germany; Ken, Kenya; SAf, South Africa; Sco, Scotland). Bootstrap values ≥70% out of 1000 replicates are shown next to the branches.

A



B

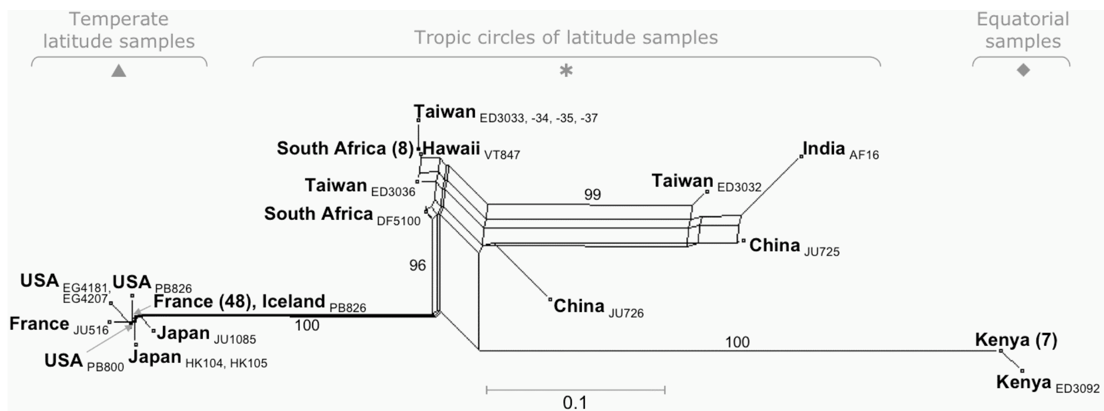


Figure 5.4: Unrooted *p*-distance neighbor-joining tree (A) and neighbor-net (B) for *C. briggsae* based on concatenated multilocus haplotypes. Clades are labeled following the latitudinal clades with the symbols in parentheses corresponding to those used to represent the strains in Figure 5.1. Bootstrap values ≥70% out of 1000 replicates are shown next to the branches.

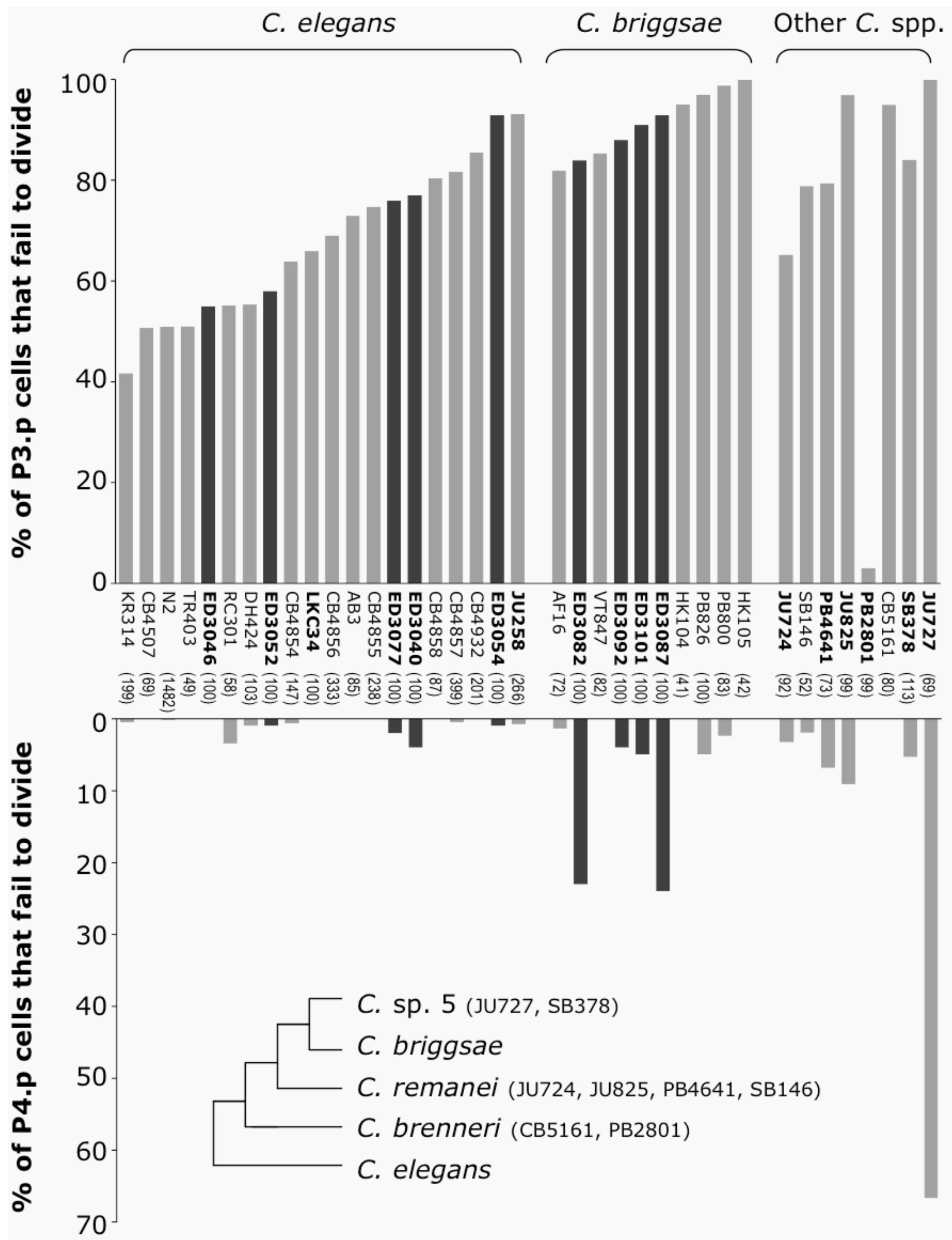


Figure 5.5: Percentage of P3.p (A) and P4.p (B) non-division in hermaphrodites of *C. elegans*, *C. briggsae*, and three other *Caenorhabditis* species of the *Elegans* group of the genus. The strain names in bold denote strains not previously analyzed in DELATTRE and FÉLIX (2001b). The darker shaded bars highlight the results for the newly collected African isolates. The number of observed worms is shown next to the strain name. The phylogenetic topology of *Caenorhabditis* species follows KIONTKE and FITCH (2005) with *C.* sp. 5 added according to K. KIONTKE and D. FITCH (personal communication).

fate, the P4.p cell was previously thought to be largely invariant in the *Caenorhabditis* genus (DELATTRE and FÉLIX 2001b); however, we found significant variation in *C. briggsae* strains for both the P3.p (χ^2_9 = 33.0, *P* = 0.0001) and P4.p cells (χ^2_9 = 94.2, *P* < 0.0001). In the two South African *C. briggsae* strains tested (ED3082 and ED3087), P4.p failed to divide 23–24% of the time. This level was significantly greater than for any other *C. briggsae* strain (all pairwise tests, $P < 10^{-8}$). In our screens of three other species in the *Elegans* group of the genus, we observed this level of P4.p non-division only in *C. briggsae*'s closest relative, *C.* sp. 5 (JU727, 67% P4.p non-division frequency, *n* = 69; SB378, 5% P4.p non-division frequency, *n* = 113). The unusual P4.p cell-lineage features found in the South African *C. briggsae* strains highlights the potential for significant phenotypic differences still to be found in these isolates, despite a high sequence similarity with the northern tropical circles of latitude strains.

5.3.3 Hybrid Crosses

Despite large differences in the molecular sequences of *C. briggsae* isolates from different parts of the world, we did not find any gross hybrid incompatibilities between African strains and strains from elsewhere, although there is substantial quantitative variation between strains for brood size and embryonic lethality (Figure 5.6). Comparisons of pure-strain versus hybrid F1s, however, show some evidence of mid-parent heterosis (MPH) for brood size, in which the hybrids show increased levels over the mean of the two parents. MPH was significantly positive in 3 of 8 crosses, and the mean MPH value across all the crosses was 16.7%. The F1 hybrids of one cross, ED3083 (South Africa) × ED3034 (Taiwan), even had significantly greater brood sizes than either pure strain parent, displaying best-parent heterosis (t_{30} = 2.07, *P* = 0.047), although this appears to be due to a large reduction in the embryonic lethality in the hybrids, which is quite high for both pure-strains (Figure 5.6).

We repeated the assays of brood size and embryonic lethality in pure-strain worms, and confirmed that these same strains displayed high proportions of dead

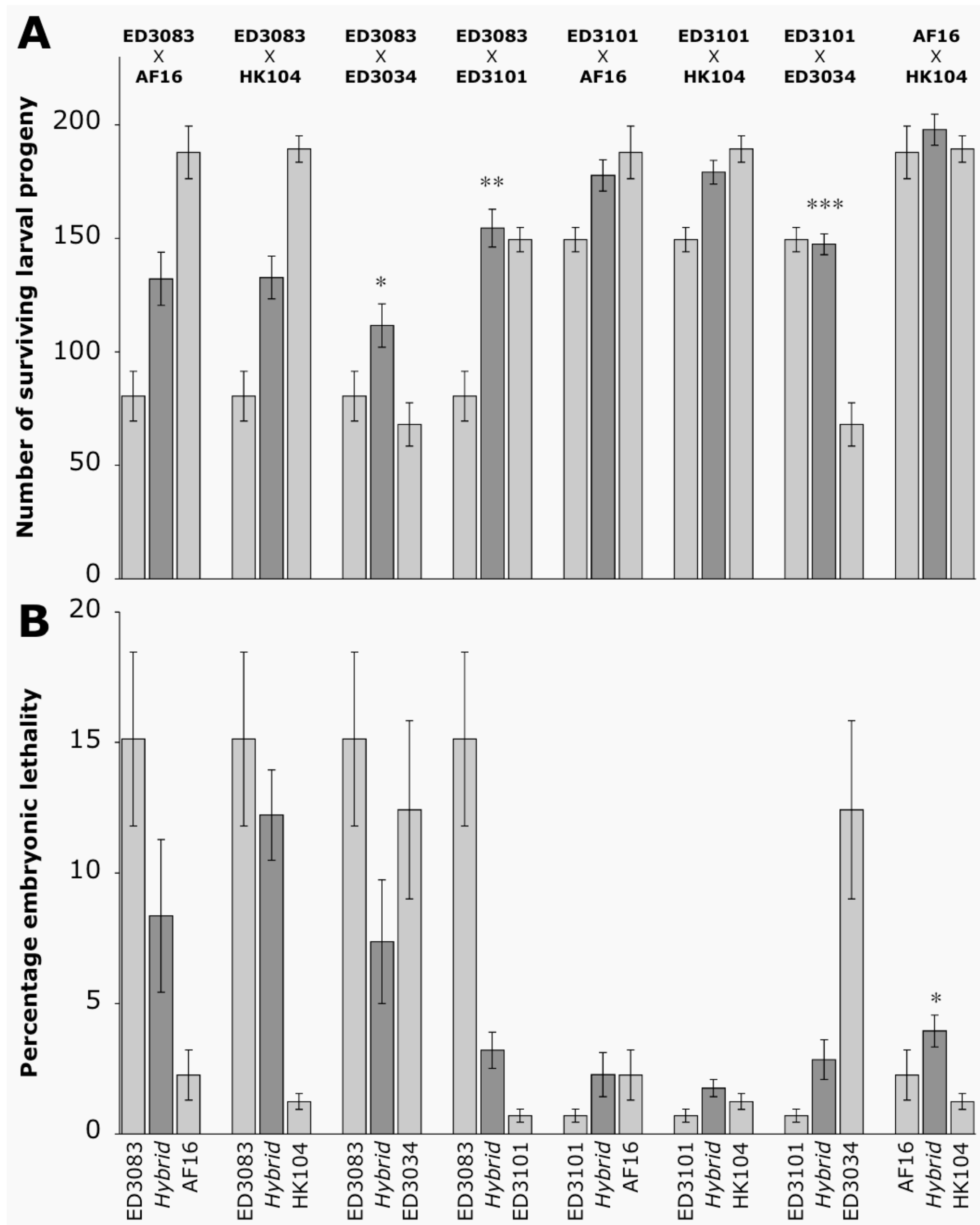


Figure 5.6: Surviving larval brood sizes (A) and embryonic lethalities (B) in the *C. briggsae* hybrid crosses. Pure-strains (lightly shaded bars) are shown next to the hybrids (darkly shaded bars) in each cross. For ease of comparison, identical data for pure-strains is shown multiple times. Asterisks indicate significance in a *t*-test of the deviation between hybrid and mid-parent. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Error bars indicate ± 1 SE.

embryos and reduced fecundities (results not shown). The low brood sizes in these strains was also independently observed elsewhere (D. DENVER, personal communication). In these strains (and others from Johannesburg, South Africa), we also observed that the worms suffered from pale and vacuolated intestines (Figure 5.7). This phenotype could result from an intracellular pathogen of intestinal cells, yet it was resistant to bleaching with alkaline hypochlorite—a method that is known to kill many extracellular pathogens, and eliminated an intracellular bacterium previously described in *C. elegans* (BARRIÈRE and FÉLIX 2005). This raises the possibility of a genetic basis to these traits, and more work is needed to explain the abnormal intestines and poor performance of these worms.

We also monitored the duration of development from egg to L4–adult molt in the pure-strain and hybrid F2s, but only observed large numbers of delayed F2s in one class of worms—the hybrids of AF16 (India) and HK104 (Japan). There was a significant difference between the duration of development of AF16 × HK104 hybrids and pure-strains ($\chi_1^2 = 33.7$, $P < 0.0001$), with ~20% of hybrid F2s exhibiting delayed development. This confirms previous observations (S. BAIRD, unpublished results), and suggests that the effect is peculiar to these genotypes, since this phenomenon was not found in any other crosses either involving the African isolates (results not shown) or other crosses without AF16 (S. BAIRD, personal communication).

5.4 Discussion

The discovery of *C. elegans* and *C. briggsae* in sub-Saharan Africa, as well as additional strains from around the world, reinforces the notion that these are indeed cosmopolitan species and offers some novel insights about levels and patterns of global variation. In *C. elegans*, this study confirms the low levels of overall silent-site polymorphism previously found in Europe for another part of the world (CUTTER 2006). In *C. briggsae,* Kenyan samples raise global species diversity estimates as a consequence of between population differences, despite a lack of variation within each of the two sampling sites in Africa. This result is similar to the only other

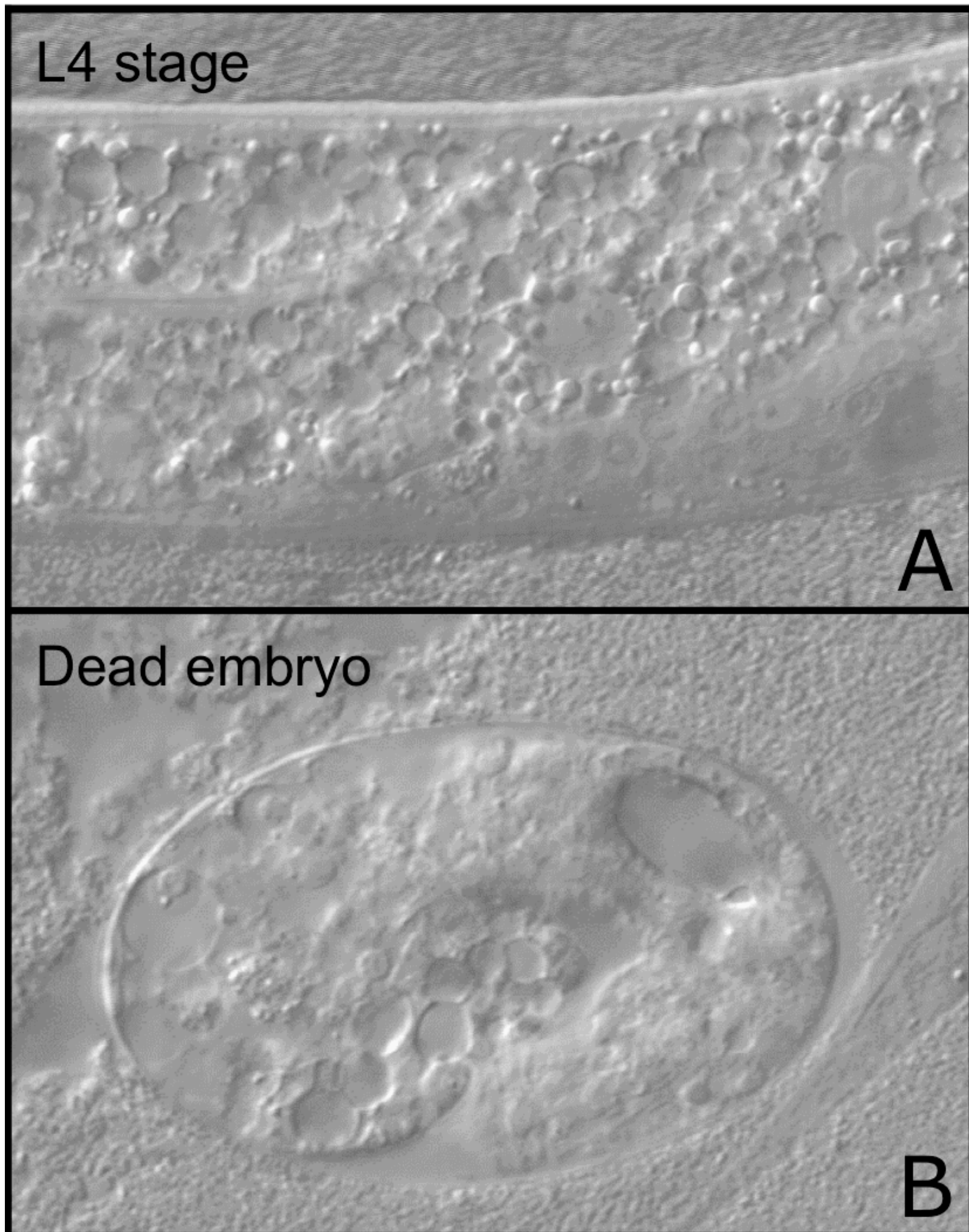


Figure 5.7: Nomarski micrographs of a South African *C. briggsae* (ED3083) L4 stage individual showing a vacuolated intestine (A) and of a dead embryo also showing large vacuoles (B). Similar phenotypes were observed for all the other *C. briggsae* strains from Johannesburg, South Africa.

known study that has established multiple wild strains from single locations (CUTTER *et al.* 2006b), although whether this reflects true levels of within-population variation in *C. briggsae* or a biased local sampling procedure remains to be seen. The natural habitat of *C. elegans* and *C. briggsae* is unknown, although both species are routinely found in anthropogenic habitats, such as compost and garden soil, and they often form phoretic and/or necromenic associations with other invertebrates (KIONTKE and SUDHAUS 2006). Given their apparent similarity in sampled habitats, it is striking that the two species have such different geographic patterning of molecular diversity (see Figures 5.3 and 5.4), even more so considering that *C. elegans* and *C. briggsae* were found co-occurring in the same sampling sites in Kenya and South Africa, as well as France (BARRIÈRE and FÉLIX 2005) and Japan (JU1085 and JU1088, described in this study). This begs the question of what could have led to the different biogeographic structures in the two species.

Population genetic studies in *C. elegans* have demonstrated that migration occurs over surprisingly large distances, including continents (BARRIÈRE and FÉLIX 2005, 2007; HABER *et al.* 2005; CUTTER 2006). Since the latitudinal clades we observe appear circumnavigatory, this may be true of *C. briggsae* as well. It has been proposed that dispersion in these species might be facilitated by human activities (KIONTKE and SUDHAUS 2006), but it is conceivable that the migration patterns and the role of human-mediated dispersion in the two species are quite different. Although both species co-occur in some localities, there may be substantial differences in their natural ecology that we have failed to appreciate. Another possibility is that *C. briggsae* more recently colonized different parts of the world than *C. elegans* (CUTTER *et al.* 2006b). However, our evidence of a third distinct geographic haplotype group in equatorial Kenya with many unique polymorphisms makes it difficult to determine the source of such a founder event. Furthermore, the lack of additional recombination events detected from the inclusion of Kenyan isolates implies that these populations might be isolated relative to other known samples. In *C. elegans*, two strains in particular (CB4856/Hawaii and JU258/Madeira), both from remote island locations, have been shown to be highly divergent from most other strains (KOCH *et al.* 2000; HABER *et al.* 2005; STEWART *et*

al. 2005; although across the six loci tested here, we did not observe this for JU258, see Figures 5.2A and 5.3), but it seems unlikely that samples from urbanized areas of Africa will be as isolated as these island strains. The absence of detectable recombination between the three clades in *C. briggsae* suggests that migration patterns and outcrossing rates could indeed be quite different in *C. briggsae* and *C. elegans*. This is further indicated by the lack of within-locality diversity in Kenya, South Africa and France (CUTTER *et al.* 2006b), which implies that gene flow is not sufficiently prevalent to prevent structuring by genetic drift and/or local selection, and is qualitatively consistent with a lower rate of outcrossing in *C. briggsae* (CUTTER *et al.* 2006b).

Our evidence for three distinct latitudinal haplotype groups in *C. briggsae*, including a "Tropic circles of latitude" clade that spans both hemispheres, suggests that ecological factors might also play a role in driving the observed patterns of biogeographic structure. One might speculate that this geographical partitioning of haplotypes reflects local adaptation. One possible outcome of this scenario would be outbreeding depression. However, we observed marginal inbreeding depression in the hybrid crosses between different latitudinal clades. In fact, outbreeding depression was previously observed for *C. elegans*, although this was probably due to the effects of the hybrid breakdown of coadapted gene complexes, rather than to effects of local adaptation (DOLGIN *et al.* 2007). Thus, it may be the case that local inbreeding and genetic drift are sufficient to explain the inbreeding depression observed between disparate population samples of *C. briggsae*. A possible methodological explanation for the inbreeding depression is that we performed our hybrid crosses under standard laboratory conditions of 20ºC, and inbreeding depression was primarily found for crosses involving ED3101 from Nairobi, Kenya. Considering that the average monthly maximum ambient temperature in Nairobi is never this low, these conditions might constitute a stressful environment for this strain, resulting in inbreeding depression (ARMBRUSTER and REED 2005), whereas experimentation under more benign conditions could produce different results.

Even though *C. elegans* displays low levels of molecular variation, wild isolates show phenotypic variation in many potentially ecologically relevant traits, including dauer sensitivity (VINEY *et al.* 2003), response to pathogens

(SCHULENBURG and MÜLLER 2004), clumping behaviour (HODGKIN and DONIACH 1997; DE BONO and BARGMANN 1998), and locomotory speed (DE BONO and BARGMANN 1998). Relating these phenotypes to their ecological importance, however, remains elusive. The low degree of variation and extensive linkage disequilibrium found in *C. elegans* make it difficult to detect signatures of local adaptation using traditional molecular population genetics approaches. The same is true for *C. briggsae*, but the more structured differences in genetic variation among geographic regions in *C. briggsae* might allow laboratory approaches, such as quantitative trait loci (QTL) mapping, to identify regions of the genome that are associated with phenotypic differences between strains from different parts of the world. In *C. elegans*, QTLs have been identified for a number of life-history related traits, with evidence of genotype–environment interactions, including in response to temperature (SHOOK and JOHNSON 1999; KNIGHT *et al.* 2001; AYYADEVARA *et al.* 2003; GUTTELING *et al.* 2007). However, the lack of correlation between biogeographic structure and diversity in *C. elegans* means it is difficult to make inferences about environmental adaptations, and suggests that any QTL identified might not necessarily correlate neatly with ecology.

The biogeographic patterning of *C. briggsae* could make this species more amenable than *C. elegans* to studies of ecological and behavioural adaptation (BAIRD and CHAMBERLIN 2006; CUTTER *et al.* 2006b). The few studies that examined strain-specific differences in *C. briggsae* have shown that strains from different clades exhibit variation in hybrid compatibilities (BAIRD 2002), the patterning of sensory rays in the male tail (BAIRD 2001), and vulval cell lineages (DELATTRE and FÉLIX 2001b). In this study, we demonstrated that additional phenotypic variation is also found within clades, as South African *C. briggsae* strains exhibited a unique pattern of P4.p division despite quite similar molecular sequences to strains from the Tropic of Cancer. It remains to be seen, however, how distinct the genomes of this particular set of strains are, and how much within- and between-clade diversity exists in the species. More samples will also be needed to determine if the latitudinal clade pattern found here is a common global feature, and to more accurately determine how the patterns of migration, outcrossing, and selection differ between *C. elegans* and *C. briggsae*. With a growing number of strains being collected from around the

world, and increasing genetic and molecular resources, *C. briggsae* presents itself as a useful companion species to *C. elegans*, and might be well-suited to studying some questions of gene flow, speciation, and adaptation.

6 Transposon Population Frequencies in *Caenorhabditis*

Contributing authors:

- I isolated DNA for the *C. elegans* samples, carried out the bioinformatics and molecular analyses, performed the statistical analyses, and wrote the manuscript.
- B. CHARLESWORTH advised on the project.
- A. CUTTER advised on the project, isolated DNA for the *C. remanei* samples, and helped write the manuscript.

6.1 Introduction

The selfish DNA hypothesis proposes that the abundance of transposable elements (TEs) in natural populations of their hosts is the consequence of a balance between the proliferation of elements by transposition and natural selection acting to remove insertions from the population (DOOLITTLE and SAPIENZA 1980; ORGEL and CRICK 1980). Accordingly, any factor that modulates the strength of selection or rates of transposition should have an important effect on the distribution and dynamics of element insertions. At the genomic level, the impact of differences in the efficacy of selection on the patterning of TEs can be seen in the genomes of many organisms, and elements have been found to accumulate differentially among chromosomal regions due to recombination rate differences (CHARLESWORTH and LANGLEY 1989; DURET *et al.* 2000; BOISSINOT *et al.* 2001; BARTOLOMÉ *et al.* 2002; RIZZON *et al.* 2002, 2003), base composition patterns (LANDER *et al.* 2001), and gene density (MEDSTRAND *et al.* 2002).

At the population level, a potentially important factor for TE dynamics is the host breeding system (HICKEY 1982; CHARLESWORTH and CHARLESWORTH 1995; WRIGHT and SCHOEN 1999; MORGAN 2001). When populations do not mate randomly, elements can be affected in a number of ways with different expected consequences. Self-fertilization can lead to element accumulation, because inbreeding reduces the effective population size experienced by the host, and smaller populations tend to accumulate more elements (CHARLESWORTH and CHARLESWORTH 1983; BROOKFIELD and BADGE 1997; WRIGHT and SCHOEN 1999; MORGAN 2001). In an inbreeding population, insertions will also tend to be homozygous, and if the deleterious effects of TEs arise because of ectopic recombination between dispersed and heterozygous insertions (MONTGOMERY *et al.* 1991), selection may be more effective in outcrossing populations. But selfing can also constrain TEs. If individual insertions are harmful because they disrupt genes (FINNEGAN 1992) and are largely recessive, inbreeding should enhance the efficacy of natural selection against TEs because of higher homozygosity (CHARLESWORTH and CHARLESWORTH 1995) and because of greater variation in copy number among individuals due to the lower rate of element transmission between individuals (HICKEY 1982; BURT and TRIVERS 1998). In addition, the complete stochastic loss of elements in self-fertilizing populations is also possible (WRIGHT and SCHOEN 1999; MORGAN 2001). A final effect of selfing is that lower levels of genetic exchange in inbred hosts might influence the evolution of transposition rates themselves, leading to conditions favouring TEs that regulate their own activity levels (CHARLESWORTH and LANGLEY 1986; CHARLESWORTH and CHARLESWORTH 1995).

Comparisons of the population frequencies of TEs in natural populations have been performed within and between many animal species, including fruitflies (CHARLESWORTH and LANGLEY 1989; PETROV *et al.* 2003; BARTOLOMÉ and MASIDE 2004; FRANCHINI *et al.* 2004), mosquitoes (O'BROCHTA *et al.* 2006; BOULESTEIX *et al.* 2007), midges (ZAMPICININI *et al.* 2004), several fish species (TAKASAKI *et al.* 1997; DUVERNELL and TURNER 1999; NEAFSEY *et al.* 2004), and humans (BATZER and DEININGER 2002; BENNETT *et al.* 2004). These studies have shed light on the selective constraints experienced by TEs, but, to date, the only empirical investigations of the effects of breeding systems on TE dynamics have focused on

closely related self- and cross-pollinating plant species (WRIGHT *et al.* 2001; TAM *et al.* 2007). Both of these studies found a significant effect of the breeding system—insertion polymorphism levels were lower and individual elements segregated at higher frequencies in self-fertilizing species. Whether these conclusions apply to TEs in selfing populations more generally remains to be seen.

In this study, we test the role of breeding systems in driving TE dynamics in two species of the nematode genus *Caenorhabditis*: the model organism *C. elegans*, which reproduces primarily as a self-fertilizing hermaphrodite, and the related outcrossing species, *C. remanei*. TEs make up about 12% of the *C. elegans* genome, with the most active and best characterized being type II DNA transposons of the Tc1/*mariner* superfamily (see review by BESSEREAU 2006). These elements, which are among the most widespread DNA transposons (PLASTERK *et al.* 1999), were named after the superfamily's two best studied members: Tc1, the first TE identified in *C. elegans* (EMMONS *et al.* 1983; LIAO *et al.* 1983), and *mariner*, which was first described in *Drosophila mauritiana* (JACOBSON *et al.* 1986). These transposons have highly diverged primary sequences, but all Tc1/*mariner* elements probably derive from a common ancestor, and they share many common features including flanking terminal inverted repeats (TIRs) and similar modes of transposition (PLASTERK *et al.* 1999). The genome of the canonical *C. elegans* strain N2 contains 32 copies of Tc1 (FISCHER *et al.* 2003), but the copy number is strain specific, and different wild strains harbour different numbers of Tc1 transposons (EMMONS *et al.* 1983; LIAO *et al.* 1983; EGILMEZ *et al.* 1995; HODGKIN and DONIACH 1997). Although these elements do not show germ line transposition in N2, they can be activated in the germ line by mutation of a single gene in the N2 strain (COLLINS *et al.* 1987; KETTING *et al.* 1999), activity can arise spontaneously in a non-mutator strain (BABITY *et al.* 1990), and some natural isolates show high rates of germ line activity (EMMONS *et al.* 1983; EIDE and ANDERSON 1985).

The *C. remanei* genome is much more poorly characterized than that of *C. elegans*, and, as such, much less is known about its TEs. Nonetheless, sequences homologous to Tc1 transposons are widely distributed among nematodes (MOERMAN and WATERSTON 1989), polymorphic Tc1-like elements have been found segregating in other nematode species (HOEKSTRA *et al.* 1999), and probes designed for the *C.*

elegans Tc1, Tc2, and Tc3 elements hybridize weakly in *C. remanei* (ABAD *et al.* 1991). In addition, the associated species *C. briggsae*, which is more closely related to *C. remanei* than *C. elegans*, contains transposons similar to Tc1, and these show polymorphic hybridization patterns (HARRIS *et al.* 1990). Here, we describe a class of Tc1-like transposons in *C. remanei*, and examine both *C. elegans* and *C. remanei* for transposon population frequencies—the degree to which individual transposons segregate in natural populations. We show that levels of insertion polymorphism significantly differ between the two species and evaluate the role of mating system in driving TE dynamics.

6.2 Materials and Methods

6.2.1 Transposable elements

The canonical N2 strain of *C. elegans* contains 32 copies of the Tc1 DNA transposon (FISCHER *et al.* 2003). We quantified the population frequencies for each of these Tc1 elements in a global sample of *C. elegans* strains. Transposon locations were taken from the genomic positions identified by FISCHER *et al.* (2003). For *C. remanei*, we identified a homologous class of transposons by using RepeatMasker (http://www.repeatmasker.org) and the Tc1 and *C. briggsae* Tcb1 sequences as the basis for *Blastn* queries of the *C. remanei* genome draft 15.0.1 for strain PB4641, which was sequenced using plasmid and fosmid libraries to a depth of 9.2X and assembled using the PCAP whole-genome assembly program (Washington University Genome Sequencing Center). The transposons we describe are similar to Tc1 and Tcb1 in the terminal regions, but they are smaller and may not be autonomous, so we have called them mTcre1 elements, where "m" stands for miniature (N. JIANG, personal communication) and "re" is an abbreviation for *C. remanei*. We determined the consensus sequence for mTcre1 transposons, and measured the population frequencies for a random subset of "full length" elements.

6.2.2 Nematode Populations and Molecular Methods

For *C. elegans,* we tested for the presence or absence of all 32 of the Tc1 elements found in the N2 strain for each of 39 *C. elegans* strains (Table 6.1). Thirty-one strains were chosen to provide worldwide representation of geographic locations where this species has been found, and an additional eight strains were selected from a single sampling locality in Hermanville, France. Genetic diversity levels at nuclear loci in Hermanville samples are similar to those found on a global level (BARRIÈRE and FÉLIX 2005; CUTTER 2006). For *C. remanei*, we quantified the presence or absence for 16 elements identified in PB4641 for the inbred strain SB146 and for 14 strains from Wright State University Biological Preserve in Dayton, Ohio, USA (Table 6.1). We selected the strains from a single locality because *C. remanei* does not show significant deviation from panmixis based on patterns of nucleotide polymorphism, and so provides a representative population sample for the species (CUTTER *et al*. 2006a).

For *C. elegans*, genomic DNA was isolated from whole plates of worms with the Puregene DNA Purification Kit (D-7000A, Gentra Systems). For *C. remanei*, we performed whole-genome amplification of single males using the REPLI-g Mini kit (Qiagen). Diluted aliquots of the DNA samples were then used as templates for standard PCR reactions. We used DNA from the sequenced strains of *C. elegans* (N2) and *C. remanei* (PB4641) as positive controls in our PCR amplification of Tc1 and mTcre1 elements, respectively.

6.2.3 Population Frequency Assays

We determined the presence or absence of individual transposon insertions in all strains using pairs of PCR primers in the flanking regions on either side of the transposon insertion sites, coupled with a third primer matching an internal portion of the transposon sequence (e.g. BARTOLOMÉ and MASIDE 2004). The PCRs performed with flanking primers produce unique sized bands that show whether the transposon is present or absent, and the reactions with the internal primers serve as additional tests to confirm transposon presence. For *C. elegans*, a single internal

Table 6.1: *C. elegans* and *C. remanei* strains used in this study.

	Strains	Location (nearest city)	Source
C. elegans	N2	Bristol, England	CGC
	AB1, AB4	Adelaide, Australia	CGC
	KR314	Vancouver, Canada	CGC
	PX174	Portland, Oregon, USA	CGC
	PS2025, CB4854	Altadena, California, USA	CGC
	CB4855	Palo Alto, California, USA	CGC
	CB4857	Claremont, California, USA	CGC
	TR403	Madison, Wisconsin, USA	CGC
	CB4856	Hawaii, USA	CGC
	JU1088	Kakegawa, Japan	MAF
	JU1171	Concepcion, Chile	MAF
	LKC34	Madagascar	MAF
	ED3040	Johannesburg, South Africa	ED
	ED3046, ED3052	Ceres, South Africa	ED
	ED3057	Limuru, Kenya	ED
	ED3077	Nairobi, Kenya	ED
	JU258	Ribeiro Frio, Madeira, Portugal	MAF
	JU776, JU780, JU799	Lisbon, Portugal	MAF
	RC301	Freiburg, Germany	CGC
	MY1, MY2, MY6, MY12	Münster, Germany	CGC
	RW7000	Bergerac, France	CGC
	JU314, JU323	Merlet, France	MAF
	JU365	Franconville, France	MAF
	JU393, JU394, JU395, JU398, JU400, JU402, JU407, JU439	Hermanville, France	MAF
C. remanei	PB4641	Brooklyn, New York, USA	CGC
	SB146	Freiburg, Germany	CGC
	PB207, PB210, PB211, PB213, PB215, PB219, PB242, PB243, PB244, PB247, PB249, PB252, PB253, PB275	Dayton, Ohio, USA	SB

CGC, Caenorhabditis Genetics Center; MAF, Marie-Anne Félix; ED, Elie Dolgin; SB, Scott Baird

primer was used for a conserved region in 31 of the 32 Tc1 insertion sites, while a different internal primer was used for clone C50H2, due to a 701-bp deletion in the transposon (FISCHER *et al.* 2003). For *C. remanei*, PCR failure rates were higher, presumably due to greater levels of nucleotide polymorphism, so we used two sets of flanking primers and unique internal primers for most transposon insertion sites. We used different combinations of primers until every strain was successfully amplified. Primer sets were designed for 21 mTcre1 insertions in PB4641, but for 5 of these, only the positive control (PB4641) worked for any primer combination. Consequently, we restrict our analysis to the 16 mTcre1 insertion sites that yielded interpretable data. The primer sequences are provided in Table 6.2.

Table 6.2: PCR primers, locations and expected fragment sizes.

	Clone or supercontig	Chrom.	Position	Primer-F	Primer-R	Internal-R	Fragment Size (bp) F&R (TE)	F&R (no TE)	F&int-R
C. elegans	F59C6	I	10,528,229	TTCTGAACCAACTAAGCCCATT	AGCCTGAGGTTTCATTCACATT	AGTCCTGCTTGCTGTAAACGTC	2015	404	811
	T27F6	I	12,481,316	TTTCCCCAAACTACACAAATCC	AACTGCATGAAATGCCTGTG	AGTCCTGCTTGCTGTAAACGTC	1938	361	983
	ZC334	I	14,021,863	GAGTTTCCTTCCGTCCAATGTA	CCACAGATTATATGCGTCAAGC	AGTCCTGCTTGCTGTAAACGTC	2030	419	815
	Y39F10A	II	784,389	ACCAAAAAGGCACCGTTAAATA	TCGTTGTAGGGGCTGTAAAGAT	AGTCCTGCTTGCTGTAAACGTC	1930	318	850
	T07D3	II	894,217	TGCTACTTGCGTGTTACATTCC	CGAGCTGAAGCAAGTATTTGTG	AGTCCTGCTTGCTGTAAACGTC	2050	439	916
	ZK250	II	1,935,482	CTGATTGTCGATTTCGAATCCT	AATTCAAGGCCAAGTTCTCAAA	AGTCCTGCTTGCTGTAAACGTC	1995	383	793
	R03H10	II	4,168,673	TGATTTTCAGGGGTTTTATTGG	CAGCACCAGGAATCATTGAATA	AGTCCTGCTTGCTGTAAACGTC	1380	451	811
	F18C5	II	6,571,170	CCTCCAATGTCGGTTATGAAAT	GTCCCTGAACCGGATAAATACA	AGTCCTGCTTGCTGTAAACGTC	2036	425	843
	C28F5	II	7,516,600	ATTGCGTGGGAAGGTAGAGATA	AGGAGTCGATGAAGCGAGTTAC	AGTCCTGCTTGCTGTAAACGTC	1941	330	802
	T21B4	II	12,520,658	CGGGCAACTTTTGTTAAACTTC	AAGAGCATGGGGGAAAATAAAT	AGTCCTGCTTGCTGTAAACGTC	1937	326	862
	Y46G5A	II	12,750,713	AAAAAGTCGAGATTTTGGCACA	GGTGAATGTGGGGAAATTACTG	AGTCCTGCTTGCTGTAAACGTC	2040	429	766
	Y48C3A	II	13,307,514	TTTGAAACGCCTCTAAAATGGT	TGGGATTTGGATTGGAAAATAG	AGTCCTGCTTGCTGTAAACGTC	1983	425	976
	Y48G9A	III	2,119,181	TATGCAACGAAATTGGATTGTC	AGAAATGTCGGCTAACGGATAA	AGTCCTGCTTGCTGTAAACGTC	1997	386	751
	Y37D8A	III	12,865,046	CGACTTGTCCAGCTTCTTTTCT	ACCTCCATGTGGAATTTTGTTC	AGTCCTGCTTGCTGTAAACGTC	2131	520	889
	Y69A2AR	IV	2,485,117	TCGATTCTACCTTGACATGTGG	CAACAACTCTGTCGAAGAATGC	AGTCCTGCTTGCTGTAAACGTC	1908	298	851
	ZK1251	IV	9,685,240	TAATTTTTGGAACGCCCATTAG	TTGGAATGTGATCTTGATTTGC	AGTCCTGCTTGCTGTAAACGTC	1958	347	836
	C48D1	IV	13,223,686	CACAAAAAGTGCTCCGTGTAAG	TGGCAAGACCTAATAAGGCAAT	AGTCCTGCTTGCTGTAAACGTC	1957	346	812
	Y73F48A	IV	15,484,060	TCTTGAACGCAAAAATGAATTG	ATTTGAAAAACGAACGGAAATG	AGTCCTGCTTGCTGTAAACGTC	1887	276	775
	T10B5	V	1,834,328	TATGCACACGGAAAAATTCAAG	ATGTGCGTTCTCATCTTCGTAA	AGTCCTGCTTGCTGTAAACGTC	1882	360	883
	T10B5	V	1,836,071	CACTTTATTGGACTTCGCCATT	CTGTGACTACGATTGTGGCTCT	AGTCCTGCTTGCTGTAAACGTC	1989	378	749
	T28A11	V	3,251,323	AGTGGTTTTCGTTTCAAGGATG	AAAATTTGGAGCCATTCAAAAA	AGTCCTGCTTGCTGTAAACGTC	2028	417	969
	T22F3	V	3,590,233	TAACCATGCGAAATCAGTTGAG	GCTGGTCTTCGAGGTTTGTACT	AGTCCTGCTTGCTGTAAACGTC	2096	484	911
	B0213	V	3,989,832	TTTCAGTGCCATATGGTGGATA	ATGTTGTGAAAATGCTGTGAGG	AGTCCTGCTTGCTGTAAACGTC	2053	442	933
	C50H2	V	9,890,306	CTAACTTGACTGCTGACGTTGG	GAGAACGGAGAGAGAAGATCCA	ATACTTTGTCGCGTATCCGTTT	1258	347	781
	ZK856	V	10,187,627	ACGTACCCCTATCGGAATTTTT	CGTGCACTTTTGGTCACATTAT	AGTCCTGCTTGCTGTAAACGTC	1851	241	759
	C03E10	V	11,288,604	CAAGCGAAAAATAAAACCGAAT	TTCATATCCCTCAGTCCCTGAT	AGTCCTGCTTGCTGTAAACGTC	2055	390	829
	F35E8	V	15,907,114	TCCAGCTTTTGACAACTTCTGA	CATCCATTGAATTTTCGGATTT	AGTCCTGCTTGCTGTAAACGTC	1962	351	901
	C31A11	V	16,312,000	TACGACAAAGGTGCTTACTGGA	TTCTTGAAGCGTTATGCGTTTA	AGTCCTGCTTGCTGTAAACGTC	1839	228	829
	Y94A7B	V	17,805,766	CACTTTTTCTCGGCTTTAATGG	TACTTTTTGGCACCTCGGTACT	AGTCCTGCTTGCTGTAAACGTC	2059	448	817
	Y51A2C	V	18,442,309	AAGCGCATAAAAGTCCTCACTC	AACTCGAGCGCTTTAAGTTACG	AGTCCTGCTTGCTGTAAACGTC	1964	353	797
	R173	X	7,017,187	GCAATATGAAAGTCGGAAGGAC	AACCATTGACAATATGCGGTCT	AGTCCTGCTTGCTGTAAACGTC	1911	300	776
	F08G12	X	11,311,126	ATCAATAGCCGCAGAGGTTAAA	GCGAATTTTCTTCAATGAAACC	AGTCCTGCTTGCTGTAAACGTC	1927	316	791
C. remanei	238	I	110,063	TTCGCTAGCCTAGAAATTCGAC	TTAATGTCGAGAGTCGGTTGTG	AAAGTCGCTCCTTTTAGCACTG	525	864	409
				ACTTCCCCCATTTTGAGTTTTT	CGATTATACGGGCAATACATCA		505	980	525
	39	II	304,491	CTGCGCTCAAACTTGTGTAATC	GCCACTCTCCTTACTAGCGAAA	CTAGCACTGAAAATGGCAACAA	568	847	396
				AATCAGCGTTTTCTCAGCTTTC	ATGCATCTTCTCACACAGCATT		263	981	530
	71	II	93,557	TCGTTGTTGTCTCCTTCTGAAA	GGGTTTGAGCATCAACAGTACA	AAATTCTCTCCTTTTGGCACTG	529	851	399
				GGTGGTAGAAATAGTGGCTTGG	AAATAATCGAATGCCTCTTGGA		393	1043	591
	111	II	200,237	GTTTCCATTATCAACCGGACAT	CAGGGCTTTGTCCTAAAATGAC	CAATCGTCAAAGTAAACAGCTCA	596	988	536
				TCTTTGTACGGTGACATTTTGG	TTAGTCAATGTGATGGCGTTTC		415	972	520
	55	III	79,358	TGAATATTAGCCGACATGAACG	ACGAGAGATGGAAATTGGAAGA	TTTGTACGGCACATCCTAAGAA	473	803	361
				GAAAATTGTCAATCGAGCCTTC	TGTCTTGCTCATGAGTTGGTCT		820	1018	576
	80	III	221,901	GTTCAGGGTTTCAGTGGACTTC	TGTGTGATTTCCGATTAGCAAC	TTTTATCACCGAAAAAGGCAAC	503	951	504
				ATTCAGAAACCACCTAGCTCCA	AGACCCAAATTATCACCACGTC		429	806	359
	237	IV	110,698	TGTTAGCCACTGAGAAGGTCAA	GGCGTGATCTATGTGTCTGAAA	TACTGTAAAACGGCAACAATCG	520	907	455
				TGAAAAACGAGACAAGAGCGTA	TTTATACTCACCGTGGCACAAG		311	997	545
	776	IV	1,344	ACGTTCAAAACATGCTTTCCTT	TCGGAGTTGTCGAAGGATTAGT	ACACGGCTTCCTTGATAGTCTC	844	1017	562
				AAGCGGAAGGGGATATAGAAAG	GAAGAAGGGAGACTCAGAACCA		433	981	526
	920	IV	8,848	ACCTCCACATCCGACTACAGTT	CACATACGGTTTGCGTTTAAGA	CAATCGTCAAAGTAAACAGCTCA	475	992	537
				AGTTGGTGTGGCACCTGTAAC	GCCGATAATAGCAATGATCACA		457	992	537
	951	IV or V	2,882	GAGAGTGAGAGACGGGAAAAGA	CACTGGAGGTCAGGTAAAGTCC	TGCGACACTTGCTGTTTTTAGT	467	907	455
				GCAAAGTGGTAAAAGCCTGAAT	TGAGAAAAGAAGCCGGATAAAG		574	976	524
	2	V	1,185,313	CATTACCCGTATCGGAAATCAT	AACTTGTTGACAGCTTCCTGGT	CAATTGCGTGAAATTTCTTTGA	428	944	492
				TCCCTGTTGTTTACGCCTTTAT	GATTGAACCATTACGGTGTTCC		550	916	464
	21	V	763,291	TGATCTCTGCTTGAGACACGAT	CTACGCAATTTCCTCAAAAACC	CAATCGTCAAAGTAAACAGCTCA	519	857	408
				TTCAAACTCGGCATGTTATGTC	GCAATTTCCTCAAAAACCACTC		592	926	477
	37	V	470,720	AATACCAGATTTTCCCGGAAGT	AAAAACGTCAAGCTTCATCCTG	TTTTAGCACTGAAAAGGGCAAT	692	979	526
				ACTAATGTTTCGCCGATGTAAA	TATGTGCATTACCGCTCAAAAG		439	1264	811
	54	V	377,140	TCGAATGTGTTTTGAAATGAGC	ATGCTTTAATTCCGGCTGAGTA	CGGGCACATCCTAAGAAGATAC	595	914	465
				GGAAATCGGAGAGAAGTGAGTG	AGAAAAATCCTGTGATGCCTTC		652	948	499
	294	V	41,101	GATGAAGACGACGACATTGGTA	TGCGGATATGTCATGAGTTAGG	CGTGTCACTGTCACAAAGTTGA	397	900	447
				TTTTTGACGACACAAAAAGTGG	ATTCGAAATCCTTCCAAACAGA		N/A	873	420
	27	X	9,685	AATTCAAAGTTTCGGCCACTTA	CTTTGCTTGCCACAGTACACTC	CTTTTGTGTACAGCCACATCCT	675	950	523
				CGGATTTCATCTCTTCCTTTTG	ACCTTTCCATTGCACAAAGATT		351	960	533

Due to high levels of somatic transposition (EMMONS and YESNER 1984), we could not differentiate homozygotes and heterozygotes when an insertion was detected. As a result, transposon presence functioned as a dominant marker. For *C. elegans*, we calculated the population frequencies by assuming that all transposons are homozygous, owing to the high degree of selfing. Therefore, the frequency of each transposon is simply equal to the number of strains for which the insertion was detected, divided by the total number of strains (excluding N2). For *C. remanei*, we first discriminated X-linked vs. autosomal insertions by inferring the chromosomal location of each insertion using *Blastn* of the unique flanking regions against the 2005 preliminary assembly of the *C. remanei* genome in Wormbase (http://www.wormbase.org). Based on WABA alignments of the resulting *C. remanei* contigs to the *C. elegans* genome, we inferred the likely syntenic

chromosome. Because we used single male DNA preparations for *C. remanei*, heterozygosity of X-linked insertions is not possible, so the frequency of such mTcre1 insertions is simply the observed population frequency. For autosomal insertions, we assumed random mating, which appears to be appropriate for the Ohio sample (CUTTER *et al*. 2006a), and calculated the frequency of each transposon according to Hardy-Weinberg expectations as one minus the square root of the frequency of strains that lacked the insertion. Frequency calculations were restricted to the 14 Ohio samples, with the inbred SB146 strain excluded from the analysis.

6.2.4 Estimating Strength of Selection

We used the diffusion approximation methods derived by PETROV *et al.* (2003) to estimate the probability that an insertion is at a particular population frequency, and calculated a maximum likelihood estimate of the selection coefficient acting on the transposons. We calculated 95% confidence intervals around the maximum likelihood scores to obtain measures of the intensity of natural selection, $N_e s$, assuming semi-dominance (i.e. h = 0.5), effective population sizes of 10^4 and 10^6 for *C. elegans* and *C. remanei*, respectively (CUTTER 2006; CUTTER *et al.* 2006a), and that all transposons have independent effects and are subject to the same strength of selection, s. Qualitative conclusions were unaffected by increasing or decreasing the effective population sizes one order of magnitude (results not shown). This analysis only considers segregating elements (i.e. not fixed), and accounts for the fact that by studying only insertions present in the reference sequenced strain, we have pre-sampled transposons in proportion to their population frequencies. For *C. elegans*, the analysis was done for the entire collection of 39 strains, for the 8 strains from Hermanville on their own, and for 30 random "scattered sample" sub-sets with a single strain selected from each geographic location (23 strains in total). This sub-sampling approach of taking a single strain from each locality may approximate a homogeneously mixing population for a large number of localities connected by migration (WAKELEY 2003; CUTTER 2006; MATSEN and WAKELEY 2006). For *C. remanei*, the analysis was limited to the Ohio population.

This analysis assumes a large number of independently segregating insertions at transposition–selection balance. To test this, we measured the level of neutral nucleotide diversity among Tc1 transposons within the genome of the canonical N2 strain and total nucleotide diversity among "full-length" mTcre1 transposons within the PB4641 *C. remanei* genome using DnaSP 4.10.9 (ROZAS *et al.* 2003). At copy number equilibrium, population genetic theory predicts that the level of neutral nucleotide diversity among active transposons is equal to $4N_e\Lambda\mu$, where μ is the neutral mutation rate, Λ is the average number of active transposons per haploid genome, and $2N_e\Lambda$ is the effective population size of the transposon family (BROOKFIELD 1986; SÁNCHEZ-GARCIA *et al.* 2005). Therefore, silent site diversity between transposons is expected to equal the average haploid copy number multiplied by the average genomic silent-site diversity, assuming that the mutation rate in transposons is similar to that of the worm's genome as a whole (BROOKFIELD 1986). We also measured levels of linkage disequilibrium among transposons for the 39 *C. elegans* strains (excluding N2) and for the 14 Ohio *C. remanei* strains, using the squared correlation between pairs of sites (r^2). Because our sample size, n, for each species is small, the effective population size multiplied by the recombination rate is expected to be much greater than the sample size, and the expected linkage disequilibrium, $E(r^2) \approx 1/n$ (WEIR and HILL 1980). Therefore, we subtract $1/n$ from our r^2-values, and compare deviations from expectations for the two species.

With *C. elegans'* well-annotated genome, we also tested for correlations between the population frequency of Tc1 transposons and various aspects of the genomic environment: local recombination rate, gene density, transposon polymorphism levels (SNPs and indels), and whether the transposon is inserted in an intron. Recombination rates based on the nearest 10 loci were taken from CUTTER and PAYSEUR (2003), numbers of SNPs and indels were obtained from FISCHER *et al.* (2003), and gene density estimates and whether the transposon was found in an intron were determined using www.wormbase.org.

6.3 Results

6.3.1 *C. remanei* Transposons

We identified a class of transposons, which we refer to as mTcre1 transposons, exhibiting high sequence similarity to the ends of the terminal inverted repeats (TIRs) of Tc1 and Tcb1 elements. Figure 6.1 shows the consensus mTcre1 sequence, which matches 17 of the outer 18-bp in the TIR of Tc1, and 30 of the outer 31-bp in the Tcb1 sequence. The mTcre1 elements have longer TIRs and a shorter total length than either Tc1 or Tcb1. The unique 86-bp internal sequence is probably too short to contain an open reading frame; however, the conserved portion of the TIRs probably contains a transposase binding site, since the Tc1 transposase binding site was identified within the outer portion of the Tc1 inverted repeats (VOS and PLASTERK 1994). In this way, mTcre1 elements might more closely resemble other Tc-family transposons such as *C. elegans* Tc7 elements, which are shorter than Tc1 and rely on Tc1-derived transposase activity (REZSOHAZY *et al.* 1997). Furthermore, mTcre1 elements are flanked by TA dinucleotides, just like Tc1 and Tcb1, which presumably result from target site sequence duplication upon integration (VAN LUENEN *et al.* 1994). We identified 81 mTcre1 transposons in total inferred from similarity to the TIRs, although this is probably an underestimate of the true number of mTcre1 elements in the PB4641 genome. Of these, 80% are "full-length" (i.e., similar in length to the consensus mTcre1 sequence), 11% contain large (>100bp) deletions, and 9% have complex insertions making them at least 400bp larger.

6.3.2 Transposon Population Frequencies

Whether each transposon was present or absent in all the strains tested is shown in Figure 6.2. *C. elegans* strains had significantly greater transposon frequencies than *C. remanei* (Mann-Whitney $U = 568.5$, $P < 10^{-6}$); on average, ~80% of the 32 Tc1 transposons found in N2 were present in any given wild strain of *C. elegans*, while *C. remanei* strains only had around half of the mTcre1 elements investigated from PB4641. The population frequencies of transposons were also

A

```
   1
   |
ta cagtactggccataaagaatgcgacacttgcagtttttagttgaaaa  47
tggtcaactttgtgactgtgacacggcctccctgatagtctcacaatatt  97
gattgtttatggagtgtatagatcacaaatagagcttcatttttgtagtt  147
gacaactttttttgtacgggcacatcctaagaagtta tcaatcgtcaaag  196
aaatttctcgcaattgttgcccttttttcagtgctaaaaggagagattttt  246
gagctgtttacattgactgcctg taacttcttaggatgtgcccgtacaa  295
aaaagttgtcaactacaaaaaatgaagctctatttgtgatctatacactcc  345
ataaacaatcaatattgtgagactatcagggaggccgtgtcacactcaca  395
aagttgaccattttcaactaaaaactgcaagtgtcgcattctttatggcc  445
agtactg ta
      |
     452
```

B

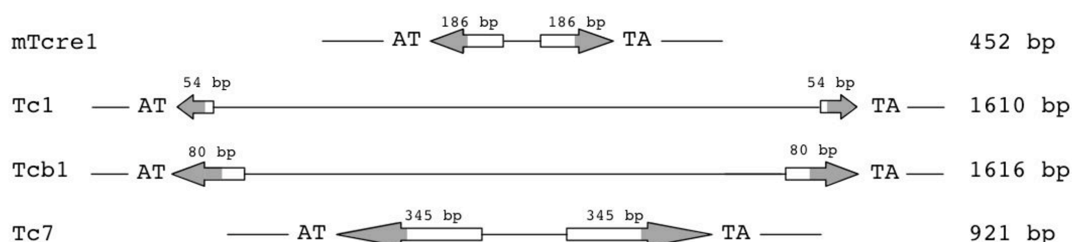
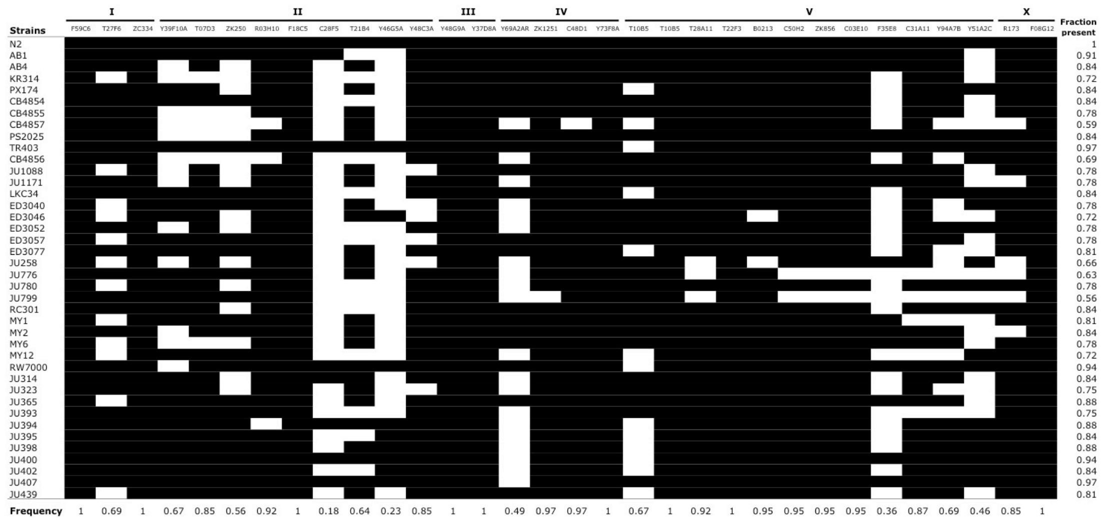


Figure 6.1: Consensus nucleotide sequence of the mTcre1 *C. remanei* transposon, with the terminal inverted repeats (TIRs) shown in boxes (A), and schematic representation of mTcre1 showing homology to Tc1, Tcb1, and Tc7 (B). Arrows represent the TIRs, with the shaded areas corresponding to regions of sequence similarity.

much higher for Tc1 elements in *C. elegans* than for mTcre1 elements in *C. remanei* (Mann-Whitney U = 454.5, $P < 10^{-6}$), as the frequency spectrum was skewed towards high frequency elements in *C. elegans* (Figure 6.3). A greater proportion of insertions were fixed for Tc1 (9 of 32 elements) than for mTcre1 (1 of 16), although this difference is not statistically significant (Fisher's Exact Test, P = 0.13). For *C. elegans*, the population frequency of Tc1 transposons did not correlate with recombination rate, gene density, polymorphism levels or whether the transposon is found in an intron (all Spearman's $|\rho| < 0.20$, $P > 0.30$). The presence/absence status

A



B

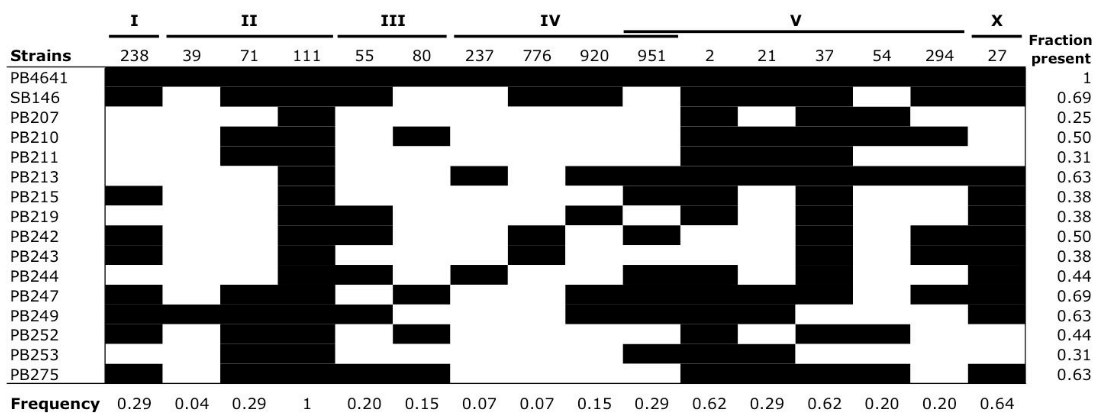


Figure 6.2: Transposon frequencies for *C. elegans* (A) and *C. remanei* (B). Dark boxes indicate presence and white boxes indicate absence of the transposon. Population frequencies of each transposon calculated for the 39 *C. elegans* strains (excluding N2), and for the 14 Ohio *C. remanei* strains are shown along the bottom; the fraction of elements found for each strain is shown along the right; and, the position of each transposon, denoted by the clone (*C. elegans*) or supercontig (*C. remanei*) and the linkage group, is shown along the top. For *C. remanei* supercontig 951, the WABA alignment was inconclusive between linkage groups IV and V.

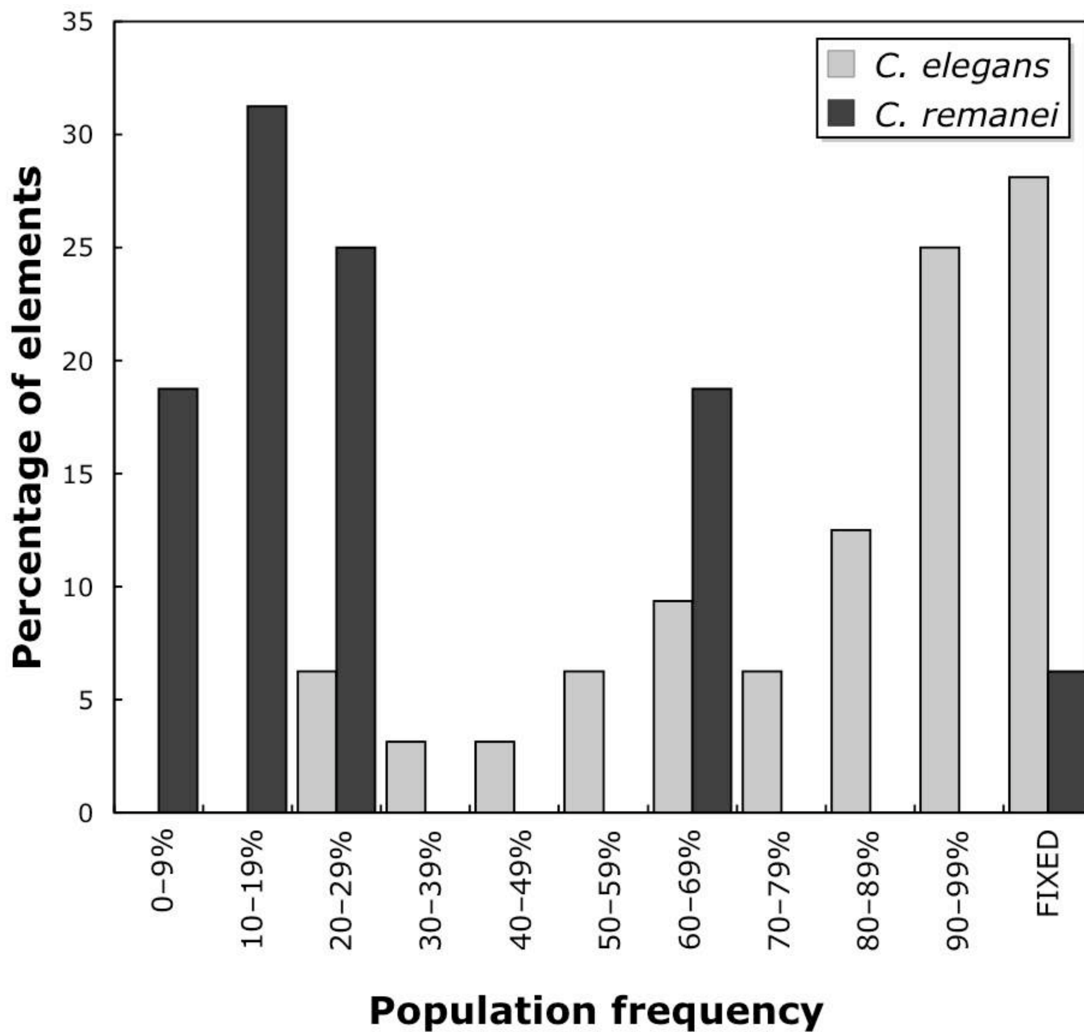


Figure 6.3: Histogram of population frequencies of *C. elegans* Tc1 transposons (light shaded bars) and *C. remanei* mTcre1 transposons (dark shaded bars).

of transposons also shows no strong signature of geographic structuring in *C. elegans*, consistent with other studies of genetic diversity (DENVER *et al.* 2003; BARRIÈRE and FÉLIX 2005; HABER *et al.* 2005; CUTTER 2006; DOLGIN *et al.* 2008).

6.3.3 Strength of Selection

Using the observed population frequency distributions of polymorphic transposons to estimate the maximum likelihood strength of selection, we found evidence for purifying selection ($N_e s < 0$) in *C. remanei* but not in *C. elegans*. For

mTcre1 elements in *C. remanei*, $N_e s$ = –2.2 (95% confidence interval: $-3.9 \leq N_e s \leq -0.6$). In contrast, when considering all *C. elegans* strains tested for Tc1 elements, $N_e s$ = 5.9 ($1.3 \leq N_e s \leq 27.3$), suggesting that Tc1 elements are subject to positive selection in *C. elegans*. This seems unlikely, however, because Tc1 transposons cause large mutagenic effects when transpositionally active in the germline (PLASTERK and VAN LUENEN 1997; BÉGIN and SCHOEN 2006, 2007).

The positive $N_e s$ value for *C. elegans* arises from the large number of high frequency transposons (see Figure 6.3), which could be an artifact of population structure if many strains shared a recent N2-like common ancestor. To attempt to remove the effects of population subdivision, we created random subsets of "scattered samples" with only a single strain from each geographic locality (WAKELEY 2003; CUTTER 2006; MATSEN and WAKELEY 2006). This analysis gave an average strength of selection, $N_e s \approx 4$. However, heterogeneity between strains from the same location led 95% confidence intervals for different scattered random subsets to imply no significant difference from neutrality in 7 of 30 subsets, and positive selection in the remaining 23. Overall, the range of lower 95% confidence intervals for the 30 random scattered subsamples spanned –0.2 to 1.1. Considering just the eight strains from a single locality in Hermanville, France, we found no significant difference from neutrality: $N_e s$ = 1.5 ($-1.1 \leq N_e s \leq 16.8$). The lack of evidence for purifying selection in *C. elegans* is also not an artifact of the assumption of complete homozygosity. If we assume Hardy-Weinberg ratios of allele frequencies, as we did for *C. remanei*, we still observe no significant departure from neutral expectations when considering all 39 strains: $N_e s$ = 0.3 ($-1.1 \leq N_e s \leq 2.8$).

The maximum likelihood method for calculating the strength of selection assumes that transposons are at copy number equilibrium. To evaluate this assumption, we contrasted predictions of nucleotide diversity and linkage disequilibrium with observed values (see Methods). For *C. remanei*, total nucleotide diversity among "full-length" mTcre1 elements in the PB4641 genome was π = 18.2% and θ = 21.3%. These diversity measures include parts of the transposon that might experience selection, including the inverted repeats and conserved region, so silent site diversity is presumably even greater. Nonetheless, these values are much greater than the 4.7% overall silent-site diversity for *C. remanei* (CUTTER *et al.*

2006a), consistent with expectations (BROOKFIELD 1986). In contrast, we found that nucleotide diversity at silent sites among *C. elegans* Tc1 elements in the N2 genome was π_{si} = 0.19% and θ_{si} = 0.76%. These values are similar to the overall silent-site diversity of 0.2–0.3% (CUTTER 2006), and not ~30X greater as predicted by equilibrium theory (BROOKFIELD 1986). In addition, Tajima's $D = -1.89$ ($P < 0.05$), indicating an excess of low-frequency variants among Tc1 elements. Furthermore, the mean deviation of linkage disequilibrium from expectation was greater for *C. elegans*: 0.056 between all sites, and 0.112 for intrachromosomal insertions, compared to r^2-deviations of 0.025 and 0.010, respectively, for *C. remanei*. A greater proportion of r^2-values were also significant by Fisher's exact tests in *C. elegans* than *C. remanei* (Fisher's exact test between species, $P = 0.002$). Together, these measures all suggest that the assumption of copy number equilibrium might be violated for *C. elegans.*

6.4 Discussion

Phylogenetic evidence indicates that the ancestor of *C. elegans* was obligately outcrossing with separate male and female individuals (KIONTKE and FITCH 2005). Thus, *C. remanei* provides a useful proxy for *C. elegans*' ancestral state, and comparisons between the two species provide a way of detecting changes in the evolutionary pressures experienced by *C. elegans* upon adopting a self-fertilizing breeding system. Analyses of sequence polymorphism show that *C. elegans* has much lower levels of genetic variation than *C. remanei* (GRAUSTEIN *et al.* 2002; JOVELIN *et al.* 2003; HAAG and ACKERMAN 2005; CUTTER 2006; CUTTER *et al.* 2006a). *C. remanei* also suffers from much stronger inbreeding depression than *C. elegans* (DOLGIN *et al.* 2007), and the two species display very different mating behaviours (CHASNOV *et al.* 2007; GARCIA *et al.* 2007). Here we show that natural populations of *C. elegans* and *C. remanei* also show markedly different patterns of transposon frequency distributions. The greater proportion of polymorphic insertions in *C. remanei* segregating at lower population frequencies compared with *C. elegans* is consistent with an important role of breeding systems contributing to the control of

transposon dynamics in natural populations, and supports other comparative studies of TEs in the plant genera *Arabidopsis* (WRIGHT *et al.* 2001) and *Solanum* (TAM *et al.* 2007).

The estimates of $N_e s$ reveal distinctly different intensities of natural selection against transposons in *C. elegans* and *C. remanei*. The negative values for *C. remanei* indicate the action of purifying selection against element insertions. On the other hand, our analyses suggest neutral or positive selection for Tc1 elements in *C. elegans*. Although TEs in general can sometimes be beneficial (KIDWELL and LISCH 2001; SCHLENKE and BEGUN 2004), this is unlikely to be generally true for Tc1 elements, which are known to cause strong deleterious mutational effects (PLASTERK and VAN LUENEN 1997; BÉGIN and SCHOEN 2006, 2007). What might bias $N_e s$ estimates for *C. elegans*? One possibility is violation of the assumption of copy number equilibrium required by the maximum likelihood method used for calculating the intensity of selection, as estimates of nucleotide diversity are lower than equilibrium predictions. However, the low estimates of genetic diversity between insertions in N2 could potentially result from gene conversion among Tc1 elements, thus biasing the molecular diversity estimates downward and also leading to a negative value for Tajima's *D*, without violating transposition–selection equilibrium. The observation that Tc1 elements can acquire the sequence of other Tc1 elements elsewhere in the genome, suggests that there might be continuous exchange of sequence information between individual insertions (FISCHER *et al.* 2003). But if gene conversion is unbiased between elements (i.e., the template transposon and the element to be converted are both random), this should not affect diversity measures (OHTA 1985; SLATKIN 1985; BROOKFIELD 1986).

The greater proportion of sites in linkage disequilibrium in *C. elegans* also suggests that the assumption of independence between insertions might be violated, although it is interesting to note that the levels of linkage disequilibrium reported here between TEs are much lower than the estimates of linkage disequilibrium in *C. elegans* from sequence data (CUTTER 2006; CUTTER *et al.* 2006a) or other molecular markers (HABER *et al.* 2005; BARRIÈRE and FÉLIX 2005, 2007). This could reflect a difference in timescale between transposons and other mutational processes. If transposition rates greatly exceed mutation rates, then transposon insertions will tend

to be more recent, making them more likely to be independent of each other. This idea is supported by the significantly negative value for Tajima's *D*, which is also consistent with a recent burst of transposition (SÁNCHEZ-GARCIA *et al.* 2005). However, this would then imply that strains lacking germ line Tc1 activity only recently acquired suppressors of transposition (COLLINS *et al.* 1987; BABITY *et al.* 1990; MORI *et al.* 1990), and the presence of numerous fixed and high frequency transposons indicates that many insertions preceded strain divergence. If we compare silent-site diversity among Tc1 elements from the N2 strain between fixed and near-fixed (≥95% population frequency) insertions with polymorphic insertions, we observe a nearly two-fold difference in diversity levels (fixed/near-fixed: π_{si} = 0.31% and θ_{si} = 0.73%; polymorphic: π_{si} = 0.17% and θ_{si} = 0.41%), further suggesting that the high frequency elements might be more ancient. Indeed, these near-fixed insertions are what drives the positive $N_e s$ values, and could simply reflect demographic structure and population subdivision in *C. elegans* (HABER *et al.* 2005; SIVASUNDAR and HEY 2005; CUTTER 2006; BARRIÈRE and FÉLIX 2005, 2007).

Frequent extinction and recolonization events or a recent population expansion also could potentially bias our results if Tc1 elements hitchhike to achieve high frequencies of elements despite their deleterious effects. If this occurs regularly on a global scale, such a scenario could affect the analyses based on a "scattered sample" of one strain from each geographic location. Indeed, metapopulation dynamics have been observed for *C. elegans* (BARRIÈRE and FÉLIX 2007), linkage disequilibrium is extensive including between chromosomes (HABER *et al.* 2005; CUTTER 2006; BARRIÈRE and FÉLIX 2005, 2007), and N2-like sequence polymorphisms are shared by many wild strains (KOCH *et al.* 2000; DENVER *et al.* 2003; BARRIÈRE and FÉLIX 2005; HABER *et al.* 2005; CUTTER 2006). Thus, the high number of fixed or near-fixed insertions across all chromosomes in *C. elegans* might reflect a genome-wide selective sweep of an N2-like genotype with low levels of excision in some strains. However, such an event would be unlikely to account for around half the insertions that show intermediate frequencies. The complete lack of low frequency (<20%) insertions in *C. elegans*, compared to around half the elements experiencing low frequencies in *C. remanei* (see Figure 6.3), suggests that the observed pattern is more likely because of selective differences between the

species. Therefore, we caution against drawing strong conclusions from the positive estimates of selection for *C. elegans* from this analysis, because of the possibility that population demographic processes might skew the transposon frequency spectrum. Rather, we argue that segregating Tc1 elements are probably selectively neutral, as seen in the analysis of the Hermanville population and some of the scattered samples.

Two main factors may be involved in causing the observed differences in transposon profiles. First, the skewed distribution in *C. elegans* towards high frequency elements suggests a reduction in the efficacy of purifying selection in selfing lineages due to a smaller effective population size (WRIGHT and SCHOEN 1999; MORGAN 2001). New transposons may persist longer in a polymorphic state in *C. remanei* and selection should be more efficient at eliminating them because its effective population size is estimated to be two orders of magnitude larger than *C. elegans*' (CUTTER 2006; CUTTER *et al.* 2006a). Assuming that transposons have similar selective effects in both species, such large population size differences could explain the different $N_e s$ estimates. But transposons could experience different selective regimes due to differences in levels of homozygosity between the breeding systems. When homozygous and heterozygous insertions have distinct fitness effects, changes in selfing rate can have dramatic effects on TE dynamics (WRIGHT and SCHOEN 1999; MORGAN 2001). Under the ectopic exchange model, selection against homozygous insertions is expected to be weak or null, whereas under the deleterious insertion model, selection will be strongest against homozygous insertions (see review by NUZHDIN 1999). Heterozygosity levels in *C. elegans* are typically very low (BARRIÈRE and FÉLIX 2005, 2007; but see SIVASUNDAR and HEY 2005), suggesting that most insertions will be in a homozygous state, which would reduce the opportunity for ectopic pairing (MONTGOMERY *et al.* 1991). Therefore, a second explanation for the high population frequencies of elements in *C. elegans* is that selection is weaker against the greater proportion of homozygous insertions, as predicted by the ectopic exchange model.

Previous studies of polymorphic Tc1 elements in *C. elegans* have compared the canonical N2 strains with only a limited number of natural isolates (EMMONS *et al.* 1983; LIAO *et al.* 1983; EIDE and ANDERSON 1985; HARRIS and ROSE 1989;

EGILMEZ *et al.* 1995; HODGKIN and DONIACH 1997). Here we studied *C. elegans* strains representing a complete global sampling covering all six major continents where *C. elegans* has been found, as well as a number of strains derived from a single geographic locality. Nevertheless, we observe many similarities in our dataset to those observed earlier for only a handful of strains. EGILMEZ *et al.* (1995) assessed Tc1-site occupancy among five strains, including N2, and found that 20 of the 32 Tc1 insertions in N2 were common to all five strains. With our much larger sample of 40 strains, we also observed high numbers of fixed insertions (9 of 32). Following the observation that the Bergerac strain had active germ line transposition and many times more Tc1 elements than N2 (EMMONS *et al.* 1983; LIAO *et al.* 1983), much of the focus on describing natural variation in Tc1 elements has been on characterizing strains as either "low copy" (~30 elements) or "high copy" (>300 elements) strains. We have not quantified copy numbers in this study, and not all the strains used have been characterized previously for transposon abundance, but it is interesting to note that the two high copy strains identified previously, RW7000/Bergerac and TR403 (EGILMEZ *et al.* 1995; HODGKIN and DONIACH 1997), were also among the strains containing the highest fraction of insertions present (see Figure 6.2A). This confirms previous findings that nearly all of the Tc1 elements in N2 are in the same location in the high copy Bergerac strain (HARRIS and ROSE 1989).

In this study, we describe a new class of Tc1-like transposons in *C. remanei*. These *C. remanei* mTcre1 elements are not strictly analogous to *C. elegans* Tc1 elements, which could potentially influence the contrast of TE polymorphism patterns between the two species. Given that the unique internal sequence of mTcre1 elements probably does not contain an open reading frame, these transposons could experience different dynamics than the autonomous Tc1 elements. Therefore, transposition rates could be quite different between Tc1 and mTcre1 elements because of innate features of the transposons. Alternatively, selection for self-regulation in selfing lineages could drive lower transposition rates in *C. elegans* (CHARLESWORTH and LANGLEY 1986), or the *effective* transposition rate might be lower in *C. elegans* if transposons cannot invade and spread into different genetic backgrounds as easily under self-fertilization. If transposition rates are indeed lower for Tc1 elements than for mTcre1 elements, fewer elements would be of recent

origin in *C. elegans*, and population frequencies could be skewed upwards. In *C. remanei*, on the other hand, if selection effectively removes most elements we might expect only to find either newly transposed elements at low frequencies or insertions that achieved high frequencies by drift. This would be consistent with the somewhat bimodal distribution for mTcre1 elements seen in Figure 6.3.

An outstanding question that cannot be answered by our method is how the total numbers of transposons in natural isolates compare between the two species. In *C. elegans*, Tc1 copy number ranges an order of magnitude between wild strains (EGILMEZ *et al.* 1995; HODGKIN and DONIACH 1997). In a single *C. remanei* genome (PB4641), we identified 81 mTcre1 elements by bioinformatics approaches, although this is likely to underestimate the true copy number, and we do not know how copy numbers vary among strains. Furthermore, it is unclear how biologically relevant it is to compare the total genomic abundance of Tc1 and mTcre1 elements if they are not homologous. Previous studies in related self- and cross-polinating plant species similarly found that outcrossing species consistently showed lower population frequencies of elements (WRIGHT *et al.* 2001; TAM *et al.* 2007), but not of element copy number. Whereas self-fertilizing *Arabidopsis* had slightly higher copy numbers of DNA transposons than outcrossing species (WRIGHT *et al.* 2001), no relationship was found between retrotransposon number and breeding system among related tomato species (TAM *et al.* 2007). Additional comparisons of TEs in other closely related species or using other TE families should help illuminate the role of breeding system on total abundance. For *Caenorhabditis*, more work is needed to characterize and quantify transposons in *C. remanei* in particular.

The results presented here support a role of breeding systems in driving TE dynamics. Two independent derivations of self-fertilization in *Caenorhabditis* (KIONTKE and FITCH 2005) among the six species slated for genome sequencing will provide a useful platform for testing the generality of breeding system as an important factor in TE evolution. Future studies will help to further tease apart the selection pressures imposed by TEs, and the impact of selfing rates on TE dynamics.

7 General discussion and conclusions

7.1 Summary

Why sexual reproduction is so widespread remains one of the most enduring questions in evolutionary biology (MAYNARD SMITH 1978; HURST and PECK 1996; BARTON and CHARLESWORTH 1998; OTTO and LENORMAND 2002). But once sex and recombination have evolved, it is equally unclear what forces promote and maintain alternative arrangements—either on the level of the whole genome, as is the case for uniparental reproduction, or at a local scale, such as when recombination is reduced in particular regions of the genome. Breeding systems and genomic recombination rates affect patterns of genetic variation within and between populations, and thus have profound effects on the rates and dynamics of evolutionary change.

One area where breeding systems should play an important role is in controlling the spread and regulation of selfish genetic elements such as TEs. When Barbara McClintock first discovered TEs in the 1940s, they were seen as an anomaly. But now we know that they make up a significant proportion of nearly all eukaryotic genomes (CRAIG *et al.* 2002). The complex relationship between TEs and their hosts had led some authors to suggest an analogy between the components of the genome, including TEs, and an ecological community (BROOKFIELD 2005; LE ROUZIC *et al.* 2007b). Under this analogy, a TE family can be seen as a species living in its habitat of the host genome. And just as a particular biological species might be well adapted to its local environment, the same could be true of TEs.

Sexual reproduction is essential for the initial spread of TEs (HICKEY 1982), and if elements have evolved to be "locally adapted" within the framework of a sexual genome, genome-wide or local reductions in the effective recombination rate can have dramatic effects on TE dynamics. Indeed, the accumulation of TEs has

been proposed as a major driving force behind the extinction of asexuals that could explain the paucity of ancient asexual eukaryotic lineages (ARKHIPOVA and MESELSON 2005a). In outcrossing populations, theoretical studies show that TEs generally evolve the highest possible transposition rates, and self-regulation of TEs is predicted to evolve only in highly selfing or asexual lineages (or when transposition happens on a very localized scale; CHARLESWORTH and LANGLEY 1986). The problem is that when an asexual lineage arises from a sexual progenitor, the TEs it contains will have evolved within a sexual genome, and these TEs will have been selected to proliferate without consideration of the ancestral sexual host's fitness. Therefore, it is by no means certain that TEs will regulate their own transposition upon the abandonment of sex in their host.

In Chapter 2, I showed that vertically transmitted TEs in asexual populations could multiply in an unbounded fashion, but only in small populations and when excision was absent. If elements do excise to some extent, analytical models predict that all deleterious TEs should be eliminated in an infinite asexual lineage, and computer simulations show this is also possible in finite populations if population sizes are large. Theoretical considerations predict that the effective population size of asexual lineages should generally be lower than that of outcrossing sexual ones, because of interference among loci in completely linked genomes (HILL and ROBERTSON 1966; FELSENSTEIN 1974; CHARLESWORTH *et al.* 1993; KEIGHTLEY and OTTO 2006), and asexuals inherit genotypes rather than alleles, thereby reducing the variance effective size (BALLOUX *et al.* 2003). Thus, most asexual lineages might not be able to escape TE-mediated extinction. But for those lineages that are large enough, the results show that complete TE elimination is possible. And if a TE-free state is achieved, this could potentially provide a long-term fitness benefit of asexuality over sexual reproduction. This relatively simple scenario of efficient selection against TEs in asexual populations with some excision could in fact account for the lack of deleterious retrotransposons in the putatively ancient asexual bdelloid rotifers (ARKHIPOVA and MESELSON 2000).

Despite decades of research on TEs in a wide variety of taxa, there is still no consensus on the way that selection acts to limit their spread. Differentiating between the different selective models has proven difficult because superficially they often

make similar predictions. With regard to recombination rates, although many authors have argued that selection is less effective against deleterious insertions in low recombination regions (e.g. DURET *et al.* 2000; BARTOLOMÉ *et al.* 2002; RIZZON *et al.* 2002), no formal modeling had previously been performed. The simulation results I presented in Chapter 3 showed that the conditions under which TEs are expected to accumulate in low recombination regions are surprisingly restrictive. By defining the parameter space allowing element accumulation under the deleterious insertion model, I created a set of testable predictions that can be compared to empirical data on the distribution and abundance of TEs. The model showed that when Hill-Robertson effects and Muller's ratchet drive the accumulation of elements in low recombination regions, insertions should generally become fixed at the population level. Currently, population frequency data is lacking, but projects such as the Drosophila Population Genomics Project (http://www.dpgp.org) should greatly improve our understanding of intraspecies natural variation and help test the predictions of the model.

The theoretical studies I presented investigated two long-standing problems concerning TE patterning in non-recombining whole or partial genomes. Turning to the lab, I explored a series of questions relating to the consequences and maintenance of androdioecy—a reproductive strategy that seems to contradict the general assumptions of breeding system evolution, because few scenarios seem to favour the maintenance of both hermaphrodites and males (LLOYD 1975; CHARLESWORTH 1984). In some ways, the results of Chapter 4 add to the enigma of this unusual breeding system. The continued presence of males in *C. elegans* has led many authors to ask if males serve some adaptive function, or whether they are simply preserved because of constant low level non-disjunction (CHASNOV and CHOW 2002; STEWART and PHILLIPS 2002; CUTTER 2003). Theory predicts that inbreeding depression experienced by selfed offspring can maintain males in an androdioecious population (LLOYD 1975; CHARLESWORTH and CHARLESWORTH 1978; OTTO *et al.* 1993), but I found little evidence for inbreeding depression in natural populations of *C. elegans*. Instead, the populations I studied showed outbreeding depression. *C. remanei*, in contrast, showed strong signs of inbreeding depression, with substantial fitness reductions in selfed progeny. These results suggest that selection might often

act against males and outcrossing in *C. elegans* in the wild. Outbreeding depression could explain why linkage disequilibrium is maintained despite a low yet significant level of outcrossing (BARRIÈRE and FÉLIX 2007), and why empirical estimates of outcrossing based on measurements reflecting different timescales often produce drastically different results (BARRIÈRE and FÉLIX 2005, 2007; SIVASUNDAR and HEY 2005; CUTTER 2006; CUTTER *et al.* 2006b).

There is strong evidence that self-fertile hermaphroditism often evolves in response to selection for reproductive assurance during colonization (PANNELL 2002; WEEKS *et al.* 2006a). The independent transitions to selfing in *C. elegans* and *C. briggsae* suggest that similar ecological pressures might have been at work. But Chapter 5 showed striking differences in the patterns of genetic and phenotypic variability in the two species. The addition of the new cultured isolates I obtained on my fieldwork in Africa adds to a growing body of work showing that the two species are more dissimilar than meets the eye (BAIRD and CHAMBERLIN 2006; GUPTA *et al.* 2007). Although reproductive assurance is probably responsible for the initial shifts to selfing in both *C. elegans* and *C. briggsae*, the two species appear to have taken different evolutionary trajectories. The results highlight the importance of conducting comparative studies between *C. elegans* and other self-fertile species to understand the evolution of androdioecy more broadly.

It is also important to compare *C. elegans* with related outcrossing species. In Chapter 6, I returned to the question of TE dynamics, and compared levels of insertion polymorphisms between *C. elegans* and *C. remanei*. To my knowledge, this was the first comparison of TE frequencies between closely related animal species with contrasting breeding systems. Elements segregated at much higher frequencies in the selfer *C. elegans* than in the outcrosser *C. remanei*, consistent with a reduction in the efficacy of natural selection under selfing. Taking the three empirical worm studies of Chapters 4–6 together, it appears that *C. elegans* is a cosmopolitan species with little correlation between genetic and geographic diversity, in which natural selection is ineffective against weakly deleterious TE insertions, and for which selection acts against crosses between strains. Given that meta-population events such as extinctions and recolonizations are probably quite common in *C. elegans* (BARRIÈRE and FÉLIX 2007), the picture that emerges is of a number of reasonably fit

independent lineages with low rates of genetic exchange occupying every continent on earth, and with an effective population size that is probably too small to eliminate many mildly deleterious mutations, including TEs. However, these are not universal features of selfing in *Caenorhabditis* nematodes, and a general description of *C. briggsae* might look quite different. It is clear though that these worms are anything but "simple multicellular organism[s]" (BRENNER 1974), and future comparative studies should further probe the implications of partial selfing for genetic and genomic architecture in these organisms.

7.2 Future directions

This thesis paves the way for many potential avenues for future research. The conclusions of Chapter 2, in which I examined the elimination of deleterious TEs in asexual populations, indicate that the complete elimination of TEs in asexual lineages depends critically on the existence of synergistic selection and occasional element excision. I proposed that the bdelloid rotifer's repair pathway for double-strand chromosomal breaks might facilitate these conditions, and thus account for their lack of deleterious retrotransposons. Studies of double-strand break repair in yeast suggest that both criteria could be met: TEs might be frequently removed by small deletions via various processes, including single-strand annealing of short homologous sequences or the removal of non-homologous 3' ends (PAQUES and HABER 1997), and synergistic selection could arise if break-induced replication causes chromosomal rearrangements (SMITH *et al.* 2007). Future empirical work could reveal whether or not the DNA repair pathway works in the same way in bdelloids. Furthermore, it would be worth testing the predictions of Chapter 2 by inspecting for TEs in the genomes of other putative ancient asexual taxa, such as the darwinulid ostracods (JUDSON and NORMARK 1996). On the theoretical front, the model I presented was one of the first formal considerations of TEs in asexual populations, but it considered only vertically-transmitted TEs, and horizontal transfer is a common feature of some classes of TEs in many organisms, including those in the bdelloids (ARKHIPOVA and MESELSON 2005b). Thus, it would be worth extending

the model to assess the impact of low rates of horizontal element transfer on the long-term fate of TEs in asexual lineages.

Chapter 3 established a theoretical framework for predicting the distribution and abundance of TEs under the deleterious insertion model when recombination rates vary across the genome. Previous studies have inspected the effects of an alternative model, the ectopic exchange model, which also predicts a relationship between recombination rate and TE abundance (LANGLEY *et al.* 1988; CHARLESWORTH *et al.* 1992a). But since these two selective mechanisms are not mutually exclusive, it would be worth formulating a unified theory to assess the outcome of both modes of selection acting together on TEs. The results of Chapter 3 also indicated a narrow parameter space where Hill-Robertson effects and Muller's ratchet could drive the accumulation of TEs under the deleterious insertion model, involving small populations, absent excision, and weak synergism between elements. Therefore, I argued that when element build-up is observed but these conditions are not met, one can generally accept the ectopic exchange model. However, when these conditions are satisfied, the ectopic exchange model could still operate, and more modeling will be needed to define rigorous criteria to differentiate between the models.

The hybrid crosses in Chapters 4 and 5 indicated that natural populations of *C. elegans* experience low levels of outbreeding depression, whereas *C. briggsae* isolates appear to suffer mild inbreeding depression. However, both studies only used a limited number of wild strains, and there was substantial variation in inbreeding depression between crosses. Thus, it would be worth expanding the number of strains tested to assess the generality of these findings. These experiments were also performed under standard laboratory conditions, and whether the results hold under more natural conditions remains to be seen. For example, variations in temperature are known to affect fecundity and reproductive timing, with significant genotype-by-environment interactions (HARVEY and VINEY 2007). Very few studies have tested *C. elegans* outside standard laboratory conditions, and fitness traits might be quite different in a more natural environment (e.g. VAN VOORHIES *et al.* 2005). In addition, other experimental tests for outbreeding depression could be used to corroborate the results. One idea would be to set up interstrain crosses and track

heterozygosity levels in the progeny through successive generations to test for selection against hybrids.

One of the most intriguing findings of Chapter 5 was that *C. briggsae* shows a much stronger signature of geographic structure than *C. elegans*. The new African strains revealed three global latitudinal clades in *C. briggsae*, but these are based on only one equatorial sampling site in Kenya and two sites near the Tropic of Capricorn in South Africa. Determining whether this pattern accurately reflects worldwide diversity will still require isolating and describing more strains from the Southern hemisphere and other "neglected" continents (e.g. South America). Even so, the results suggest that the ecology of the two species is quite different. The classical picture of *C. elegans* as a soil nematode is quickly vanishing, but we still know embarrassingly little about its natural environment (KIONTKE and SUDHAUS 2006). In Africa, I found *C. elegans* and *C. briggsae* in anthropogenic locations, but there is increasing evidence that rotting fruit might be a more natural habitat (M.-A. FÉLIX, personal communication). Systematic sampling and a better understanding of the worms' natural history will be crucial to fully understand the costs, benefits, and consequences of androdioecy in the genus.

The conclusion that selfing rates affect transposon population frequencies was consistent with previous studies in plants. But the PCR-based approach of Chapter 6 only inspected insertion polymorphism levels, and it remains to be seen how total transposon abundance compares in natural populations of *C. elegans* and *C. remanei*. Moreover, the transposons tested were not strictly homologous between the species and greater characterization of TEs in the *C. remanei* genome is needed. In future work, it would thus be worth examining the dynamics of a number of different element families. In the *C. elegans* genome, DNA-based transposons are preferentially located in high recombination regions, but no such relationship was found for retrotransposons (DURET *et al.* 2000). The patterning of TEs in the *C. remanei* genome is unknown, but these results suggest that selection may act differently on DNA- and RNA-based elements, and cross-species comparisons of population dynamics in selfing and outcrossing populations could be useful to illuminate the similarities and differences in the role of selection on different types of elements.

This thesis has tackled a variety of different problems relating to breeding system evolution, but it has also *opened a can of worms* for future research.

Bibliography

ABAD, P., C. QUILES, S. TARES, C. PIOTTE, P. CASTAGNONE-SERENO *et al.*, 1991 Sequences homologous to Tc(s) transposable elements of *Caenorhabditis elegans* are widely distributed in the phylum Nematoda. *J. Mol. Evol*. **33**: 251–258.

ADAMS, M. D., S. E. CELNIKER, R. A. HOLT, C. A. EVANS, J. D. GOCAYNE *et al.*, 2000 The genome sequence of *Drosophila melanogaster*. *Science* **287**: 2185–2195.

ALLARD, R. W., 1975 Mating system and microevolution. *Genetics* **79**: 115–126.

ANTONOVICS, J., 1968 Evolution in closely adjacent populations. V. Evolution of self-fertility. *Heredity* **23**: 219–238.

ARKHIPOVA, I., and M. MESELSON, 2000 Transposable elements in sexual and ancient asexual taxa. *Proc. Natl. Acad. Sci.* USA **97**: 14473–14477.

ARKHIPOVA, I., and M. MESELSON, 2005a Deleterious transposable elements and the extinction of asexuals. *BioEssays* **27**: 76–85.

ARKHIPOVA, I., and M. MESELSON, 2005b Diverse DNA transposons in rotifers of the class Bdelloidea. *Proc. Natl. Acad. Sci.* USA **102**: 11781–11786.

ARKHIPOVA, I. R., K. I. PYATKOV, M. MESELSON and M. B. EVGEN'EV, 2003 Retroelements containing introns in diverse invertebrate taxa. *Nat. Genet.* **33**: 123–124.

ARMBRUSTER, P., and D. H. REED, 2005 Inbreeding depression in benign and stressful environments. *Heredity* **95**: 235–242.

ASHBURNER, M., K. G. GOLIC and R. S. HAWLEY, 2004 *Drosophila: A Laboratory Handbook*. 2nd edition. Cold Spring Harbor Press, Cold Spring Harbor, NY.

AYYADEVARA, S., R. AYYADEVARA, A. VERTINO, A. GALECKI, J. J. THADEN, and R. J. SHMOOKLER REIS, 2003 Genetic loci modulating fitness and life span in *Caenorhabditis elegans*: categorial trait interval mapping in CL2a x Beregerac-BO recombinant-inbred worms. *Genetics* **163**: 557–570.

BABITY, J. M., T. V. B. STARR and A. M. ROSE, 1990 Tc1 transposition and mutator activity in a Bristol strain of *Caenorhabditis elegans*. *Mol. Gen. Genet.* **222**: 606–611.

BACHTROG, D., 2003 Accumulation of *Spock* and *Worf*, two novel non-LTR retrotransposons, on the neo-Y chromosome of *Drosophila miranda*. *Mol. Biol. Evol.* **20**: 173–181.

BACHTROG, D., and I. GORDO, 2004 Adaptive evolution of asexual populations under Muller's ratchet. *Evolution* **58**: 1403–1413.

BAKER, H. G., 1955 Self-compatibility and establishment after "long-distance" dispersal. *Evolution* **9**: 347–349.

BAÏLLE D., A. BARRIÈRE, and M.-A. FÉLIX 2008 *Oscheius tipulae*, a widespread hermaphroditic soil nematode, displays a higher genetic diversity and geographic structure than *Caenorhabditis elegans*. *Mol. Ecol.*, in press.

BAIRD, S. E., 1999 Natural and experimental association of *Caenorhabditis remanei* with *Trachelipus rathkii* and other terrestrial isopods. *Nematology* **1**: 471–475.

BAIRD, S. E., 2001 Strain-specific variation in the pattern of caudal papillae in *Caenorhabditis briggsae* (Nematoda: Rhabditidae); implications for species identification. *Nematology* **3**: 373–376.

BAIRD, S. E., 2002 Haldane's rule by sexual transformation in *Caenorhabditis*. *Genetics* **161**: 1349–1353.

BAIRD, S. E., and H. M. CHAMBERLIN, 2006 *Caenorhabditis briggsae* methods (December 18, 2006), in *Wormbook*, edited by THE *C. ELEGANS* RESEARCH COMMUNITY, http://www.wormbook.org.

BAIRD, S. E., C. R. DAVIDSON, and J. C. BOHRER, 2005 The genetics of ray pattern variation in *Caenorhabditis briggsae*. *BMC Evol. Biol.* **5**: 3.

BALLOUX, F., L. LEHMANN and T. DE MEEÛS, 2003 The population genetics of clonal and partially clonal diploids. *Genetics* **164**: 1635–1644.

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

BARRIÈRE, A., and M.-A. FÉLIX, 2006 Isolation of *C. elegans* and related nematodes (July 17, 2006), in *Wormbook*, edited by THE *C. ELEGANS* RESEARCH COMMUNITY, http://www.wormbook.org.

BARRIÈRE, A., and M.-A. FÉLIX, 2007 Temporal dynamics and linkage disequilibrium in natural *Caenorhabditis elegans* populations. *Genetics* **176**: 999–1011.

BARTOLOMÉ, C., and X. MASIDE, 2004 The lack of recombination drives the fixation of transposable elements on the fourth chromosome of *Drosophila melanogaster*. *Genet. Res.* **83**: 91–100.

BARTOLOMÉ, C., X. MASIDE and B. CHARLESWORTH, 2002 On the abundance and distribution of transposable elements in the genome of Drosophila melanogaster. *Mol. Biol. Evol.* **19**: 926–937.

BARTON, N., and B. O. BENGTSSON, 1986 The barrier to genetic exchange between hybridizing populations. *Heredity* **57**: 357–376.

BARTON, N. H., and B. CHARLESWORTH, 1998 Why sex and recombination? *Science* **281**: 1986–1990.

BASTEN, C. J., and M. E. MOODY, 1991 A branching-process model for the evolution of transposable elements incorporating selection. *J. Math. Biol.* **29**: 743–761.

BATTISTA, J. R., 1997 Against all odds: The survival strategies of *Deinococcus radiodurans*. *Annu. Rev. Microbiol.* **51**: 203–224.

BATZER, M. A., and P. L. DEININGER, 2002 *Alu* repeats and human genomic diversity. *Nat. Rev. Genet.* **3**: 370–379.

BÉGIN, M., and D. J. SCHOEN, 2006 Low impact of germline transposition on the rate of mildly deleterious mutation in *Caenorhabditis elegans*. *Genetics* **174**: 2129–2136.

BÉGIN, M., and D. J. SCHOEN, 2007 Transposable elements, mutational correlations, and population divergence in *Caenorhabditis elegans*. *Evolution* **61**: 1062–1070.

BENNETT, E. A., L. E. COLEMAN, C. TSUI, W. S. PITTARD and S. E. DEVINE, 2004 Natural genetic variation caused by transposable elements in humans. *Genetics* **168**: 933–951.

BERGMAN, C. M., and D. BENSASSON, 2007 Recent LTR retrotransposon insertion contrasts with waves of non-LTR insertion since speciation in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci.* USA **104**: 11340–11345.

BERGMAN, C. M., H. QUESNEVILLE, D. ANXOLABÉHÈRE and M. ASHBURNER, 2006 Recurrent insertion and duplication generate networks of transposable element sequences in the *Drosophila melanogaster* genome. *Genome Biol.* **7**: R112.

BESSEREAU, J.-L., 2006 Transposons (January 18, 2006), in *Wormbook*, edited by THE *C. ELEGANS* RESEARCH COMMUNITY, http://www.wormbook.org.

BIÉMONT, C., F. LEMEUNIER, M. P. GARCIA GUERREIRO, J. F. Y. BROOKFIELD, C. GAUTIER *et al.*, 1994 Population dynamics of *copia*, *mdg1*, *mdg3*, *gypsy*, and *P* transposable elements in a natural population of *Drosophila melanogaster*. *Genet. Res.* **63**: 197–212.

BIÉMONT, C., A. TSITRONE, C. VIEIRA and C. HOOGLAND, 1997 Transposable element distributions in Drosophila. *Genetics* **147**: 1997–1999.

BIRKY, C. W. Jr., C. WOLF, H. MAUGHAN, L. HERBERTSON and E. HENRY, 2005 Speciation and selection without sex. *Hydrobiologia* **546**: 29–45.

BOISSINOT, S., A. ENTEZAM and A. V. FURANO, 2001 Selection against deleterious LINE-1-containing loci in the human lineage. *Mol. Biol. Evol.* **18**: 926–935.

BOULESTEIX, M., F. SIMARD, C. ANTONIO-NKONDJIO, H. P. AWONO-AMBENE, D. FONTENILLE *et al.*, 2007 Insertion polymorphism of transposable elements and population structure of *Anopheles gambiae* M and S molecular forms in Cameroon. *Mol. Ecol.* **16**: 441–452.

BRAENDLE, C., and M.-A. FÉLIX, 2006 Sex determination: Ways to evolve a hermaphrodite. *Curr. Biol.* **16**: R468–R471.

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

BROOKFIELD, J. F. Y., 1986 A model for DNA sequence evolution within transposable element families. *Genetics* **112**: 393–407.

BROOKFIELD, J. F. Y., 1991 Models of repression of transposition in *P-M* hybrid dysgenesis by *P* cytotype and by zygotically encoded repressor proteins. *Genetics* **128**: 471–486.

BROOKFIELD, J. F. Y., 2005 The ecology of the genome: mobile DNA elements and their host. *Nat. Rev. Genet.* **6:** 128–136.

BROOKFIELD, J. F. Y., and R. M. BADGE, 1997 Population genetic models of transposable elements. *Genetica* **100**: 281–294.

BURKE, J. M., and M. L. ARNOLD, 2001 Genetics and the fitness of hybrids. *Annu. Rev. Genet.* **35**: 31–52.

BURT, A., and R. TRIVERS, 1998 Selfish DNA and breeding system in flowering plants. *Proc. R. Soc. Lond.* B **265**: 141–146.

BUTCHER, D., 1995 Muller's ratchet, epistasis and mutation effects. *Genetics* **141**: 431–437.

C. ELEGANS SEQUENCING CONSORTIUM, 1998 Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* **282**: 2012–2018.

CARR, D. E., and M. R. DUDASH, 2003 Recent approaches into the genetic basis of inbreeding depression in plants. *Phil. Trans. R. Soc. Lond.* B **358:** 1071–1084.

CARR, M., J. R. SOLOWAY, T. E. ROBINSON, and J. F. Y. BROOKFIELD, 2001 An investigation of the cause of low variability on the fourth chromosome of *Drosophila melanogaster*. Mol. Biol. Evol. **18**: 2260–2269.

CASALS, F., J. GONZÁLEZ and A. RUIZ, 2006 Abundance and chromosomal distribution of six *Drosophila buzzatii* transposons: *BuT1*, *BuT2*, *BuT3*, *BuT4*, *BuT5*, and *BuT6*. *Chromosoma* **115:** 403–412.

CHARLESWORTH, B., 1980 The cost of sex in relation to mating system. *J. Theor. Biol.* **84:** 655–671.

CHARLESWORTH, B., 1985 The population genetics of transposable elements, pp. 213-232 in *Population Genetics and Molecular Evolution*, edited by T. OHTA and K. AOKI. Springer-Verlag, Berlin.

CHARLESWORTH, B., 1990 Mutation-selection balance and the evolutionary advantage of sex and recombination. *Genet. Res.* **55**: 199–221.

CHARLESWORTH, B., 1991 Transposable elements in natural populations with a mixture of selected and neutral insertion sites. *Genet. Res.* **57**: 127–135.

CHARLESWORTH, B., 1992 Evolutionary rates in partially self-fertilizing species. *Am. Nat.* **140**: 126–148.

CHARLESWORTH, B., 1994 Evolution in Age-structured Populations. Cambridge University Press, Cambridge, UK.

CHARLESWORTH, B., 1996 Background selection and patterns of genetic diversity in *Drosophila melanogaster*. Genet. Res. **68**: 131–149.

CHARLESWORTH, B., and D. CHARLESWORTH, 1983 The population dynamics of transposable elements. *Genet. Res.* **42**: 1–27.

CHARLESWORTH, D., and B. CHARLESWORTH, 1995 Transposable elements in inbreeding and outbreeding populations. *Genetics* **140**: 415–417.

CHARLESWORTH, B., and D. CHARLESWORTH, 1998 Some evolutionary consequences of deleterious mutations. *Genetica* **102/103**: 3–19.

CHARLESWORTH, B., and D. CHARLESWORTH, 1999 The genetic basis of inbreeding depression. *Genet. Res.* **74**: 329–340.

CHARLESWORTH, B., and D. CHARLESWORTH, 2000 The degeneration of Y chromosomes. *Phil. Trans. R. Soc. Lond.* B **355**: 1563–1572.

CHARLESWORTH, B., and C. H. LANGLEY, 1986 The evolution of self-regulated transposition of transposable elements. *Genetics* **112**: 359–383.

CHARLESWORTH, B., and C. H. LANGLEY, 1989 The population genetics of *Drosophila* transposable elements. *Annu. Rev. Genet.* **23**: 251–287.

CHARLESWORTH, B., and C. H. LANGLEY, 1991 Population genetics of transposable elements in *Drosophila*, pp. 150-176 in *Evolution at the Molecular Level*, edited by R. K. SELANDER, A. G. CLARK and T. S. WHITTAM. Sinauer Associates, Sunderland, MA.

CHARLESWORTH, B., A. LAPID and D. CANADA, 1992a The distribution of transposable elements within and between chromosomes in a population of *Drosophila melanogaster*. I. Element frequencies and distribution. *Genet. Res.* **60**: 103–114.

CHARLESWORTH, B., A. LAPID and D. CANADA, 1992b The distribution of transposable elements within and between chromosomes in a population of

Drosophila melanogaster. II. Inferences on the nature of selection against elements. *Genet*. *Res*. **60**: 115–130.

CHARLESWORTH, B., M. T. MORGAN and D. CHARLESWORTH, 1993 The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**: 1289–1303.

CHARLESWORTH, B., P. SNIEGOWSKI and W. STEPHAN, 1994 The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* **371**: 215–220.

CHARLESWORTH, B., C. H. LANGLEY and P. D. SNIEGOWSKI, 1997 Transposable element distributions in *Drosophila*. *Genetics* **147**: 1993-1995.

CHARLESWORTH, D., 1984 Androdioecy and the evolution of dioecy. *Biol*. *J*. *Linn*. *Soc*. **22**: 333–348.

CHARLESWORTH, D., 2003 Effects of inbreeding on the genetic diversity of populations. *Phil. Trans. R. Soc. Lond*. B **358**: 1051–1070.

CHARLESWORTH, D., and B. CHARLESWORTH, 1995 Transposable elements in inbreeding and outbreeding populations. *Genetics* **140**: 415–417.

CHARLESWORTH, D., M. T. MORGAN and B. CHARLESWORTH, 1990 Inbreeding depression, genetic load and the evolution of outcrossing rates in a multi-locus system with no linkage. *Evolution* **44**: 1469–1489.

CHARLESWORTH, D., M. T. MORGAN, and B. CHARLESWORTH, 1993 Mutation accumulation in finite populations. *J. Hered*. **84**: 321–325.

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

CHASNOV, J. R., W. K. SO, C. M. CHAN and K. L. CHOW, 2007 The species, sex, and stage specificity of a *Caenorhabditis* sex pheromone. *Proc*. *Natl*. *Acad*. *Sci*. USA **104**: 6730–6735.

COGHLAN, A., and K. H. WOLFE, 2002 Fourfold faster rate of genome rearrangement in nematodes than in *Drosophila*. *Genome Res.* **12**: 857–867.

COLLINS, J., B. SAARI and P. ANDERSON, 1987 Activation of a transposable element in the germ line but not the soma of *Caenorhabditis elegans*. *Nature* **328**: 726–728.

COMBADÃO, J., P. R. A. CAMPOS, F. DIONISIO and I. GORDO, 2007 Small-world networks decrease the speed of Muller's ratchet. *Genet*. *Res*. **89**: 7–18.

COMERON, J. M., and M. KREITMAN, 2002 Population, evolutionary and genomic consequences of interference selection. *Genetics* **161**: 389–410.

COYNE, J. A., and H. A. ORR, 1998 The evolutionary genetics of speciation. *Phil*. *Trans. R. Soc. Lond*. B **353**: 287–305.

CRAIG, N. L., R. CRAIGIE, M. GELLERT and A. M. LAMBOWITZ (Editors), 2002 *Mobile DNA II*. ASM Press, Washington.

CRNOKRAK, P., and S. C. H. BARRETT, 2002 Purging the genetic load: A review of the experimental evidence. *Evolution* **56**: 2347–2358.

CROW, J. F., 1970 Genetic loads and the cost of natural selection in *Mathematical topics in population genetics*, edited by K.-I. KOJIMA. Springer-Verlag, Berlin.

CUTTER, A. D., 2004 Sperm-limited fecundity in nematodes: How many sperm are enough? *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., and B. A. PAYSEUR, 2003 Selection at linked sites in the partial selfer *Caenorhabditis elegans*. *Mol. Biol Evol.* **20**: 665–673.

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., L. AVILES, AND S. WARD, 2003 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., M.-A. FÉLIX, A. BARRIÈRE and D. CHARLESWORTH, 2006b Patterns of nucleotide polymorphism distinguish temperate and tropical wild isolates of *Caenorhabditis briggsae*. *Genetics* **173**: 2021–2031.

DARWIN, C. R., 1876 *The Effects of Cross and Self Fertilisation in the Vegetable Kingdom*. John Murray, London.

DE BONO, M., and C. I. BARGMANN, 1998 Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans*. *Cell* **94**: 679–689.

DELATTRE, M., and M.-A. FÉLIX, 2001a Microevolutionary studies of nematodes: a beginning. *BioEssays* **23**: 807–819.

DELATTRE, M., and M.-A. FÉLIX, 2001b Polymorphism and evolution of vulval precursor cell lineages within two nematode genera, *Caenorhabditis* and *Oscheius*. *Curr. Biol.* **11**: 631–643.

DEMEESTER, L., 1993 Inbreeding and outbreeding depression in *Daphnia*. *Oecologia* **96**: 80–84.

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.

DOCKING, T. R., F. E. SAADÉ, M. C. ELLIOTT and D. J. SCHOEN, 2006 Retrotransposon sequence variation in four asexual plant species. *J. Mol. Evol.* **62**: 375–387.

DOLGIN, E. S., and B. CHARLESWORTH, 2006 The fate of transposable elements in asexual populations. *Genetics* **174**: 817–827.

DOLGIN, E. S., B. CHARLESWORTH, S. E. BAIRD, and A. D. CUTTER, 2007 Inbreeding and outbreeding depression in *Caenorhabditis* nematodes. *Evolution* **61**: 1339–1352.

DOLGIN, E. S., M.-A. FÉLIX and A. D. CUTTER, 2008 Hakuna Nematoda: genetic and phenotyic diversity in African isolates of *Caenorhabditis elegans* and *C. briggsae*. *Heredity*, in press.

DOOLITTLE, W. F., and C. SAPIENZA, 1980 Selfish genes, the phenotype paradigm and genome evolution. *Nature* **284**: 601–603.

DOOLITTLE, W. F., T. B. L. KIRKWOOD and M. A. H. DEMPSTER, 1984 Selfish DNA and self-restraint. *Nature* **307**: 501–502.

DURET, L., G. MARAIS and C. BIÉMONT, 2000 Transposons but not retrotransposons are located preferentially in regions of high recombination rate in *Caenorhabditis elegans*. *Genetics* **156**: 1661–1669.

DUVERNELL, D. D., and B. J. TURNER, 1999 Variation and divergence of death valley pupfish populations at retrotransposon-defined loci. *Mol. Biol. Evol.* **16**: 363–371.

EANES, W. F., C. WESLEY and B. CHARLESWORTH, 1992 Accumulation of *P* elements in minority inversions in natural populations of *Drosophila melanogaster*. *Genet. Res.* **59**: 1–9.

EDWARDS, R. J., and J. F. Y. BROOKFIELD, 2003 Transiently beneficial insertions could maintain mobile DNA sequences in variable environments. *Mol. Biol. Evol.* **20**: 30–37.

EGILMEZ, N. K., R. H. EBERT, II and R. J. SHMOOKLER REIS, 1995 Strain evolution in *Caenorhabditis elegans*: transposable elements are markers of interstrain evolutionary history. *J. Mol. Evol.* **40**: 372–381.

EIDE, D., and P. ANDERSON, 1985 Transposition of Tc1 in the nematode *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* USA **82**: 1756–1760.

EMLEN, J. M., 1991 Heterosis and outbreeding depression: a multilocus model and an application to salmon production. *Fish. Res.* **12**: 187–212.

EMMONS, S. W., and L. YESNER, 1984 High-frequency excision of transposable element Tc1 in the nematode *Caenorhabditis elegans* is limited to somatic cells. *Cell* **36**: 599–605.

EMMONS, S. W., L. YESNER, K. S. RUAN and D. KATZENBERG, 1983 Evidence for a transposon in *Caenorhabditis elegans*. *Cell* **32**: 55–65.

EWENS, W. J., 2004 *Mathematical Population Genetics*. 2nd edition. Springer-Verlag, Berlin.

FALCONER, D. S., and T. F. C. MACKAY, 1996 Introduction to Quantitative Genetics. Longman Group Ltd., Essex.

FELSENSTEIN, J., 1974 The evolutionary advantage of recombination. *Genetics* **78**: 737–756.

FENSTER, C. B., and M. R. DUDASH, 1994 Genetic considerations for plant population restoration and conservation, pp. 34-62 in *Restoration of endangered species*, edited by M. L. BOWLES and C. J. WHELAN. Cambridge University Press, Cambridge, UK.

FINNEGAN, D. J., 1992 Transposable elements, pp. 1096-1107 in *The genome of Drosophila melanogaster*, edited by D. L. LINDSLEY and G. ZIMM. Academic Press, New York.

FISCHER, S. E. J., E. WIENHOLDS and R. H. A. PLASTERK, 2003 Continuous exchange of sequence information between dispersed Tc1 transposons in the *Caenorhabditis elegans* genome. *Genetics* **164**: 127–134.

FISHER, R. A., 1941 Average excess and average effect of a gene substitution. *Ann. Eugen*. **11**: 53–63.

FLOYD, R., E. ABEBE, A. PAPERT, and M. BLAXTER, 2002 Molecular barcodes for soil nematode identification. *Mol. Ecol.* **11**: 839–850.

FODOR, A., D. L. RIDDLE, F. K. NELSON, and J. W. GOLDEN, 1983 Comparison of a new wild-type *Caenorhabditis briggsae* with laboratory strains of *Caenorhabditis briggsae* and *C. elegans*. *Nematologica* **29**: 203–217.

FONTANILLAS, P., D. L. HARTL, and M. REUTER, 2007 Genome organization and gene expression shape the transposable element distribution in the *Drosophila melanogaster* euchromatin. *PLoS Genet*. **3**: e210.

FRANCHINI, L. F., E. W. GANKO and J. F. MCDONALD, 2004 Retrotransposon-gene associations are widespread among *D. menalogaster* populations. *Mol. Biol. Evol.* **21**: 1323–1331.

FULLERTON, S. M., A. BERNARDO CARVALHO and A. G. CLARK, 2001 Local rates of recombination are positively correlated with GC content in the human genome. *Mol. Biol. Evol*. **18**: 1139–1142.

GABRIEL, W., M. LYNCH and R. BURGER, 1993 Muller's Ratchet and mutational meltdowns. *Evolution* **47**: 1744–1757.

GARCIA, L. R., B. LEBOEUF and P. KOO, 2007 Diversity in mating behavior of hermaphroditic and male-female *Caenorhabditis* nematodes. *Genetics* **175**: 1761–1771.

GLÉMIN, S., 2003 How are deleterious mutations purged? Drift versus nonrandom mating. *Evolution* **57**: 2678–2687.

GLÉMIN, S., J. RONFORT, and T. BATAILLON, 2003 Patterns of inbreeding depression and architecture of the load in subdivided populations. *Genetics* **165**: 2193–2212.

GOLDMAN, A. S., and M. LICHTEN, 1996 The efficiency of meiotic recombination between dispersed sequences in *Sacchromyces cerevisiae* depends upon their chromosomal location. *Genetics* **144**: 43–55.

GOLDMAN, A. S., and M. LICHTEN, 2000 Restriction of ectopic recombination by interhomolog interactions during *Saccharomyces cerevisiae* meiosis. *Proc. Natl. Acad. Sci.* USA **97**: 9537–9542.

GOODWILLIE, C., 1999 Multiple origins of self-compatibility in *Linanthus* section Leptosiphon (Polemoniaceae): Phylogenetic evidence from internal-transcribed-spacer sequence data. *Evolution* **53**: 1387–1395.

GOODWILLIE, C., S. KALISC, and S. G. ECKERT, 2005 The evolutionary enigma of mixed mating systems in plants: Occurrence, theoretical explanations, and empirical evidence. *Annu. Rev. Ecol. Evol. Syst.* **36**: 47–79.

GORDO, I., and B. CHARLESWORTH, 2000a The degeneration of asexual haploid populations and the speed of Muller's ratchet. *Genetics* **154**: 1379–1387.

GORDO, I., and B. CHARLESWORTH, 2000b On the speed of Muller's ratchet. *Genetics* **156**: 2137–2140.

GORDO, I., and B. CHARLESWORTH, 2001 The speed of Muller's ratchet with background selection and the degeneration of Y chromosomes. *Genet. Res.* **78**: 149–161.

GORDO, I., A. NAVARRO and B. CHARLESWORTH, 2002 Muller's ratchet and the pattern of variation at a neutral locus. *Genetics* **161**: 835–848.

GRAUSTEIN, A., J. M. GASPAR, J. R. WALTERS and M. F. PALOPOLI, 2002 Levels of DNA polymorphism vary with mating system in the nematode genus *Caenorhabditis*. *Genetics* **161**: 99–107.

GUPTA, B. P., R. JOHNSEN and N. CHEN, 2007 Genomics and biology of the nematode *Caenorhabditis briggsae* (May 3, 2007), in *Wormbook*, edited by THE *C. ELEGANS* RESEARCH COMMUNITY, http://www.wormbook.org.

GUTTELING, E. W., J. A. G. RIKSEN, J. BAKKER, J. E. KAMMENGA, 2007 Mapping phenotypic plasticity and genotype-environment interactions affecting life-history traits in *Caenorhabditis elegans*. *Heredity* **98**: 28–37.

GVOZDEV, V. A., G. L. KOGAN and L. A. USAKIN, 2005 The Y chromosome as a target for acquired and amplified genetic material in evolution. *BioEssays* **27**: 1256–1262.

HAAG, E. S., and A. D. ACKERMAN, 2005 Intraspecific variation in *fem-3* and *tra-2*, two rapidly coevolving nematode sex-determining genes. *Gene* **349**: 35–42.

HAAG, E. S., H. CHAMBERLIN, A. COGHLAN, D. H. A. FITCH, A. D. PETERS *et al.*, 2007 *Caenorhabditis* evolution: if they all look alike, you aren't looking hard enough. *Trends Genet.* **23**: 101–104.

HABER, M., M. SCHÜNGEL, A. PUTZ, S. MÜLLER, B. HASERT *et al.*, 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.

HAIGH, J., 1978 The accumulation of deleterious genes in a population – Muller's ratchet. *Theor. Popul. Biol.* **14**: 251–267.

HALLIGAN, D. L., A. D. PETERS, and P. D. KEIGHTLEY, 2003 Estimating numbers of EMS-induced mutations affecting life history traits in *Caenorhabditis elegans* in crosses between inbred sublines. *Genet. Res.* **82**: 191–205.

HARRIS, L. J., S. PRASAD and A. M. ROSE, 1990 Isolation and sequence analysis of *Caenorhabditis briggsae* repetitive elements related to the *Caenorhabditis elegans* transposon Tc1. *J. Mol. Evol.* **30**: 359–369.

HARRIS, L. J., and A. M. ROSE, 1989 Structural analysis of Tc1 elements in *Caenorhabditis elegans* var. Bristol (strain N2). *Plasmid* **22**: 10–21.

HARTL, D. L., and S. A. SAWYER, 1988 Why do unrelated insertion sequences occur together in the genome of *Escherichia coli*? *Genetics* **118**: 537–541.

HARVEY, S. C., and M. E. VINEY, 2007 Thermal variation reveals natural variation between isolates of *Caenorhabditis elegans*. *J. Exp. Zoolog.* (*Mol. Dev. Evol.*) **308B**: 409–416.

HAUSER, T. P., and V. LOESCHCKE, 1995 Inbreeding depression in *Lychnis flos-cuculi* (Caryophyllaceae): Effects of different levels of inbreeding. *J. Evol. Biol.* **8**: 589–600.

HAYES, C. N., J. A. WINSOR, and A. G. STEPHENSON, 2005 A comparison of male and female responses to inbreeding in Cucurbita pepo subsp. texana (Cucurbitaceae). *Am. J. Bot.* **92**: 107–115.

HEDRICK, P. W., 1994 Purging inbreeding depression and the probability of extinction: full sib-mating. *Heredity* **73**: 363–372.

HENRY, P.-Y., R. PRADEL, and P. JARNE, 2003 Environment-dependent inbreeding depression in a hermaphroditic freshwater snail. *J. Evol. Biol.* **16**: 1211–1222.

HICKEY, D. A., 1982 Selfish DNA: a sexually-transmitted nuclear parasite. *Genetics* **101**: 519–531.

HILL, R. C., C. E. DE CARVALHO, J. SALOGIANNIS, B. SCHLAGER, D. PILGRIM *et al.*, 2006 Genetic flexibility in the convergent evolution of hermaphroditism in *Caenorhabditis* nematodes. *Dev. Cell* **10**: 531–538.

HILL, W. G., and A. ROBERTSON, 1966 The effect of linkage on limits to artificial selection. *Genet. Res.* **8**: 269–294.

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. DONIACH, 1997 Natural variation and copulatory plug formation in *Caenorhabditis elegans*. *Genetics* **146**: 149–164.

HODGKIN, J., H. R. HORVITZ and S. BRENNER, 1979 Nondisjunction mutants of the nematode *Caenorhabditis elegans*. *Genetics* **91**: 67–94.

HOEKSTRA, R., M. OTSEN, J. A. LENSTRA and M. H. ROOS, 1999 Characterisation of a polymorphic Tc1-like transposable element of the parasitic nematode *Haemonchus contortus*. *Mol. Biochem. Parasitol.* **102**: 157–166.

HOLSINGER, K. E., 2000 Reproductive systems and evolution in vascular plants. *Proc. Natl. Acad. Sci.* USA **97**: 7037–7042.

HOLTSFORD, T. P., and N. C. ELLSTRAND, 1990 Inbreeding effects in *Clarkia tembloriensis* (Onagraceae) populations with different natural outcrossing rates. *Evolution* **44**: 2031–2046.

HONDA, H., 1925 Experimental and cytological studies on bisexual and hermaphrodite free-living nematodes, with special reference to problems of sex. *J. Morph.* **40**: 191–233.

HOOGLAND, C., and C. BIÉMONT, 1996 Chromosomal distribution of transposable elements in *Drosophila melanogaster*: test of the ectopic recombination model for maintenance of insertion site number. *Genetics* **144**: 197–204.

HUA-VAN, A., A. LE ROUZIC, C. MAISONHAUTE and P. CAPY, 2005 Abundance, distribution and dynamics of retrotransposable elements and transposons: similarities and differences. *Cytogenet. Genome Res.* **110**: 426–440.

HUDSON, R. R., 2000 A new statistic for detecting differentiation. *Genetics* **155**: 2011–2014.

HUDSON, R. R., D. D. BOOS, and N. L. KAPLAN, 1992 A statistical test for detecting geographic subdivision. *Mol. Biol. Evol.* **9**: 138–151.

HUDSON, R. R., and N. L. KAPLAN, 1985 Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* **111**: 147–164.

HURST, L. D., and J. R. PECK, 1996 Recent advances in understanding of the evolution and maintenance of sex. *Trends Ecol. Evol.* **11**: 46–52.

HUSON, D. H., and D. BRYANT, 2006 Applications of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**: 254–267.

HUSBAND, B. C., and D. W. SCHEMSKE, 1996 Evolution of the magnitude and timing of inbreeding depression in plants. *Evolution* **50**: 54–70.

INOUE, T., M. AILION, S. POON, H. K. KIM, J. H. THOMAS, P. W. STERNBERG, 2007 Genetic analysis of dauer formation in *Caenorhabditis briggsae*. *Genetics* **177**: 809–818.

JACOBSON, J. W., M. M. MEDHORA and D. L. HARTL, 1986 Molecular structure of a somatically unstable transposable element in *Drosophila*. *Proc. Natl. Acad. Sci.* USA **83:** 8684–8688.

JARNE, P., and D. CHARLESWORTH, 1993 The evolution of the selfing rate in functionally hermaphroditic plants and animals. *Annu. Rev. Ecol. Syst.* **24**: 441–466.

JIANG, M., J. RUY, M. KIRALY, K. DUKE, V. REINKE, and S. K. KIM, 2001 Genome-wide analysis of developmental and sex-regulated gene expression profiles in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* USA **98**: 218–223.

JOHNSON, T., and N. H. BARTON, 2002 The effect of deleterious alleles on adaptation in asexual populations. *Genetics* **162**: 395–411.

JOHNSON, T. E., and E. W. HUTCHINSON, 1993 Absence of strong heterosis for life-span and other life-history traits in *Caenorhabditis elegans*. *Genetics* **134**: 465–474.

JOHNSON, T. E., and W. B. WOOD, 1982 Genetic analysis of lifespan in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* USA **79**: 6603–6607.

JOHNSTON, M. O., and D. J. SCHOEN, 1996 Correlated evolution of self-fertilization and inbreeding depression: An experimental study of nine populations of *Amsinckia* (Boraginaceae). *Evolution* **50**: 1478–1491.

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.

JORDAN, I. K., L. V. MATYUNINA and J. F. MCDONALD, 1999 Evidence for recent horizontal transfer of long terminal repeat retrotransposons. *Proc. Natl. Acad. Sci.* USA **96**: 12621–12625.

JUDSON, O. P., and B. B. NORMARK, 1996 Ancient asexual scandals. *Trends Ecol. Evol.* **11**: 41–46.

KAMINKER, J. S., C. M. BERGMAN, B. KRONMILLER, J. CARLSON, R. SVIRSKAS *et al.*, 2002 The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. *Genome Biol.* **3**: RESEARCH0084.

KAPLAN, N. L., and J. F. Y. BROOKFIELD, 1983 Transposable elements in Mendelian populations. III. Statistical results. *Genetics* **104**: 485–495.

KEIGHTLEY, P. D., and A. CABALLERO, 1997 Genomic mutation rates for lifetime reproductive output and lifespan in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* USA **94**: 3823–3827.

KEIGHTLEY, P. D., and S. P. OTTO, 2006 Interference among deleterious mutations favours sex and recombination in finite populations. *Nature* **443**: 89–92.

KEIGHTLEY, P. D., E. K. DAVIES, A. D. PETERS, and R. G. SHAW, 2000 Properties of ethylmethane sulfonate-induced mutations affecting life-history traits in *Caenorhabditis elegans* and inferences about bivariate distributions of mutation effects. *Genetics* **156**: 143–154.

KELLER, L. F., and D. M. WALLER, 2002 Inbreeding effects in wild population. *Trends Ecol. Evol.* **17**: 230–241.

KETTING, R. F., T. H. HAVERKAMP, H. G. A. M. VAN LUENEN and R. H. A. PLASTERK, 1999 *mut-7* of *C. elegans*, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. *Cell* **99**: 133–141.

KIDWELL, M. G., 2002 Transposable elements and the evolution of genome size in eukaryotes. *Genetica* **115**: 49–-63.

KIDWELL, M. G., and D. R. LISCH, 2001 Transposable elements, parasitic DNA, and genome evolution. *Evolution* **55**: 1–24.

KIMURA, M., and T. MARUYAMA, 1966 The mutational load with epistatic gene interactions in fitness. *Genetics* **54**: 1303–1312.

KIMURA, M., and T. OHTA, 1971 Theoretical Aspects of Population Genetics. Princeton University Press, Princeton, NJ.

KIONTKE, K., and D. H. FITCH, 2005 The phylogenetic relationships of *Caenorhabditis* and other rhabditids (August 11, 2005), in *Wormbook*, edited by THE *C. ELEGANS* RESEARCH COMMUNITY, http://www.wormbook.org.

KIONTKE, K., and W. SUDHAUS, 2006 Ecology of *Caenorhabditis* species (January 9, 2006), in *Wormbook*, edited by THE *C. ELEGANS* RESEARCH COMMUNITY, http://www.wormbook.org.

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*. USA **101**: 9003–9008.

KNIGHT, T., 1799 Experiments on the fecundation of vegetables. *Phil. Trans. R. Soc. Lond*. B **89**: 195–204.

KNIGHT, C. G., R. B. R. AZEVEDO, and A. M. LEROI, 2001 Testing life-history pleiotropy in *Caenorhabditis elegans*. *Evolution* **55**: 1795–1804.

KOCH, R., H. G. A. M. VAN LUENEN, M. VAN DER HORST, K. L. Thijssen, and R. H. A. P. PLASTERK, 2000 Single nucleotide polymorphisms in wild isolates of *Caenorhabditis elegans*. *Genome Res*. **10**: 1690–1696.

KONDRASHOV, A. S., 1993 Classification of hypotheses on the advantage of amphimixis. *J. Hered*. **84:** 372–387.

KONDRASHOV, A. S., 1994 Muller's ratchet under epistatic selection. *Genetics* **136**: 1469–1473.

KREITMAN, M., 1983 Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. *Nature* **304**: 412–417.

LAKOWSKI, B., and S. HEKIMI, 1996 Determination of life-span in *Caenorhabditis elegans* by four clock genes. *Science* **272**: 1010–1013.

LANDE, R., and D. W. SHEMSKE, 1985 The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. *Evolution* **48**: 965–978.

LANDER, E. S., L. M. LINTON, B. BIRREN, C. NUSBAUM, M. C. ZODY *et al.*, 2001 Initial sequencing and analysis of the human genome. *Nature* **409**: 860–921.

LANGLEY, C. H., E. A. MONTGOMERY, R. HUDSON, N. KAPLAN and B. CHARLESWORTH, 1988 On the role of unequal exchange in the containment of transposable element copy number. *Genet. Res.* **52**: 223–236.

LAPINSKI, J., and A. TUNNACLIFFE, 2003 Anhydrobiosis without trehalose in bdelloid rotifers. *FEBS Lett.* **553**: 387–390.

LE ROUZIC, A., and G. DECELIERE, 2005 Models of the population genetics of transposable elements. *Genet. Res.* **85**: 171–181.

LE ROUZIC, A., T. S. BOUTIN and P. CAPY, 2007a Long-term evolution of transposable elements. *Proc. Natl. Acad. Sci.* USA **104**: 19375–19380.

LE ROUZIC, A., S. DUPAS and P. CAPY, 2007b Genome ecosystem and transposable elements species. *Gene* **390**: 214–220

LEWIS, H., 1973 The origin of diploid neospecies in *Clarkia*. *Am. Nat.* **107**: 161–170.

LIAO, L. W., B. ROSENZWEIG and D. HIRSH, 1983 Analysis of a transposable element in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* USA **80**: 3585–3589.

LLOYD, D., 1975 The maintenance of gynodioecy and androdioecy in angiosperms. *Genetica* **45**: 325–339.

LOEWE, L., 2006 Quantifying the genomic decay paradox due to Muller's ratchet in human mitochondrial DNA. *Genet. Res.* **87**: 133–159.

LYNCH, M., 1991 The genetic interpretation of inbreeding depression and outbreeding depression. *Evolution* **45**: 622–629.

LYNCH, M., 1996 Mutation accumulation in transfer RNAs: molecular evidence for Muller's ratchet in mitochondrial genomes. *Mol. Biol. Evol.* **13**: 209–220.

LYNCH, M., 2007 The Origins of Genome Architecture. Sinauer Associates, Sunderland, MA.

LYNCH, M., and L. BLANCHARD, 1998 Deleterious mutation accumulation in organelle genomes. *Genetica* **102/103**: 29–39.

LYNCH, M., and W. GABRIEL, 1990 Mutation load and the survival of small populations. *Evolution* **44**: 1725–1737.

LYONS, E. E., 1996 Breeding system evolution in *Leavenworthia.* II. Genetic and nongenetic parental effects on reproductive success in selfing and more outcrossing populations of *Leavenworthia crassa*. *Am. Nat.* **147**: 65–85.

MALIK, H. S., W. D. BURKE and T. H. EICKBUSH, 1999 The age and evolution of non-LTR retrotransposable elements. *Mol. Biol. Evol.* **16**: 793–805.

MANOEL, D., S. CARVALHO, P. C. PHILLIPS and H. TEOTÓNIO, 2007 Selection against males in *Caenorhabditis elegans* under two mutational treatments. *Proc. R. Soc.* B **274**: 417–424.

MARK WELCH, D., and M. MESELSON, 2000 Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. *Science* **28**: 49–50.

MARTIEL, J.-L., and M. BLOT, 2002 Transposable elements and fitness of bacteria. *Theor. Popul. Biol.* **61**: 509–518.

MASIDE, X., S. ASSIMACOPOULOS and B. CHARLESWORTH, 2000 Rates of movement of transposable elements on the second chromosome of *Drosophila melanogaster*. *Genet. Res.* **75**: 275–284.

MASIDE, X., C. BARTOLOMÉ, S. ASSIMACOPOULOS and B. CHARLESWORTH, 2001 Rates and movement of transposable elements on the second chromosome of *Drosophila melanogaster*: in situ hybridization versus Southern blotting data. *Genet. Res.* **78**: 121–136.

MASIDE, X., S. ASSIMACOPOULOS and B. CHARLESWORTH, 2005 Fixation of transposable elements in the *Drosophila melanogaster* genome. *Genet. Res.* **85**: 195–203.

MASLIN, T. P., 1968 Taxonomic problems in parthenogenetic vertebrates. *Syst. Zool.* **17**: 219–231.

MATSEN, F. A., and J. WAKELEY, 2006 Convergence to the island-model coalescent process in populations with restricted migration. *Genetics* **172**: 701–708.

MAUPAS, E. F., 1900 Modes et formes de reproduction des nématodes. *Arch. Zool. Expér. Gén.* **8**: 463–624.

MATSUMOTO, M., and T. NISHIMURA, 1998 Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator. *ACM Trans. Mod. Comput. Simul.* **8**: 3–30.

MAYDAN, J. S., S. FLIBOTTE, M. L. EDGLEY, J. LAU, R. R. SELZER, T. A. RICHMOND TA *et al.*, 2007 Efficient high-resolution deletion discovery in *Caenorhabditis elegans* by array comparative genomic hybridization. *Genome Res.* **17**: 337–347.

MAYNARD SMITH, J., 1978 *The Evolution of Sex*. Cambridge University Press, Cambridge, UK.

MAYNARD SMITH, J., 1986 Contemplating life without sex. *Nature* **324**: 300–301.

MAYNARD SMITH, J., 1992 Age and the unisexual lineage. *Nature* **356**: 661–662.

MAYNARD SMITH, J., and J. HAIGH, 1974 The hitch-hiking effect of a favorable gene. *Genet. Res.* **231**: 1114–1116.

MCVEAN, G. A. T., and B. CHARLESWORTH, 2000 The effects of Hill-Robertson interference between weakly selected mutations on patterns of molecular evolution and variation. *Genetics* **155**: 929–944.

MEDSTRAND, P., L. N. VAN DE LAGEMAAT and D. L. MAGER, 2002 Retroelement distributions in the human genome: variations associated with age and proximity to genes. *Genome Res.* **12**: 1483–1495.

MOERMAN, D. G., and R. H. WATERSTON, 1989 Mobile elements in *Caenorhabditis elegans* and other nematodes, pp. 537-556 in *Mobile DNA*, edited by D. E. BERG and M. H. HOWE. American Society for Microbiology, Washington, DC.

MONTGOMERY, E. A., B. CHARLESWORTH and C. H. LANGLEY, 1987 A test for the role of natural selection in the stabilization of transposable element copy number in a population of *Drosophila melanogaster*. *Genet. Res.* **49**: 31–41.

MONTGOMERY, E. A., S. M. HUANG, C. H. LANGLEY and B. H. JUDD, 1991 Chromosome rearrangement by ectopic recombination in *Drosophila melanogaster*: genome structure and evolution. *Genetics* **129**: 1085–1098.

MORAN, N. A., 1996 Accelerated evolution and Muller's ratchet in endosymbiont bacteria. *Proc. Natl. Acad. Sci.* USA **93**: 2873–2878.

MORGAN, M. T., 2001 Transposable element number in mixed mating populations. *Genet. Res.* **77**: 261–275.

MORI, I., D. G. MOERMAN and R. H. WATERSTON, 1990 Interstrain crosses enhance excision of Tc1 transposable elements in *Caenorhabitis elegans*. *Mol. Gen. Genet.* **220**: 251–255.

MULLER, H. J., 1964 The relation of recombination to mutational advance. *Mutat. Res.* **1**: 2–9.

NAYAK, S., J. GOREE, and T. SCHEDL, 2005 *fog-2* and the evolution of self-fertile hermaphroditism in *Caenorhabditis*. *PLoS Biol.* **3**: e6.

NEAFSEY, D. E., J. P. BLUMENSTIEL and D. L. HARTL, 2004 Different regulatory mechanisms underlie similar transposable element profiles in pufferfish and fruitflies. *Mol. Biol. Evol.* **21**: 2310–2318.

NEI, M., T. MARUYAMA and R. CHAKRABORTY, 1975 The bottleneck effect and genetic variability in populations. *Evolution* **29**: 1–10.

NIGON, V., and E. C. DOUGHERTY, 1949 Reproductive patterns and attempts at reciprocal crossing of *Rhabditis elegans* Maupas, 1900, and *Rhabditis briggsae*

Dougherty and Nigon, 1949 (Nematoda: Rhabditidae). *J. Exp. Zool.* **112**: 485–503.

NORDBORG, M., 2000 Linkage disequilibrium, gene trees and selfing: an ancestral recombination graph with partial self-fertilization. *Genetics* **154**: 923–929.

NUZHDIN, S. V., 1999 Sure facts, speculations, and open questions about evolution of transposable elements. *Genetica* **107**: 129–137.

NUZHDIN, S. V., and T. F. C. MACKAY, 1995 The genomic rate of transposable element movement in *Drosophila melanogaster*. *Mol. Biol. Evol.* **12**: 180–181.

NUZHDIN, S. V., and D. A. PETROV, 2003 Transposable elements in clonal lineages: lethal hangover from sex. *Biol. J. Linn. Soc.* **79**: 33–41.

NUZHDIN, S. V., E. G. PASYUKOVA, T. V. MOROZOVA and A. J. FLAVELL, 1998 Quantitative genetic analysis of *copia* retrotransposon activity in inbred *Drosophila melanogaster* lines. *Genetics* **150**: 755–766.

O'BROCHTA, D. A., R. A. SUBRAMANIAN, J. ORSETTI, E. PECKHAM, N. NOLAN *et al.*, 2006 *hAT* element population genetics in *Anopheles gambiae s.l.* in Mozambique. *Genetica* **127**: 185–198.

OHTA, T., 1985 A model of duplicative transposition and gene conversion for repetitive DNA families. *Genetics* **110**: 513–524.

ORGEL, L. E., and F. H. CRICK, 1980 Selfish DNA: the ultimate parasite. *Nature* **284**: 604–607.

OTTO, S. P., and T. LENORMAND, 2002 Resolving the paradox of sex and recombination. *Nat. Rev. Genet.* **3**: 252–261.

OTTO, S. P., C. SASSAMAN, and M. W. FELDMAN, 1993 Evolution of sex determination in the crustacean shrimp *Eulimnadia texana*. *Am. Nat.* **141**: 329–337.

PANNELL, J. R., 2002 The evolution and maintenance of androdioecy. *Annu. Rev. Ecol. Syst.* **33**: 397–425.

PAQUES, F., and J. E. HABER, 1997 Two pathways for removal of nonhomologous DNA ends during double-strand break repair in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **17**: 6765–6771.

PASYUKOVA, E. G., S. V. NUHZDIN and D. A. FILATOV, 1998 The relationship between the rate of transposition and transposable element copy number for copia and Doc retrotransposons of *Drosophila melanogaster*. *Genet. Res.* **72**: 1–11.

PETERS, A. D., D. L. HALLIGAN, M. C. WHITLOCK, and P. D. KEIGHTLEY. 2003. Dominance and overdominance of mildly deleterious induced mutations for fitness traits in *Caenorhabditis elegans*. *Genetics* **165**: 589–599.

PETROV, D., Y. T. AMINETZRACH, J. C. DAVIS, D. BENSASSON and A. E. HIRSH, 2003 Size matters: non-LTR retrotransposable elements and ectopic recombination in *Drosophila*. *Mol. Biol. Evol.* **20**: 880–892.

PHILLIPS, P. C., and N. A. JOHNSON, 1998 The population genetics of synthetic lethals. *Genetics* **150**: 449–458.

PLASTERK, R. H. A., Z. IZSVÁK and Z. IVICS, 1999 Resident aliens: the Tc1/*mariner* superfamily of transposable elements. *Trends Genet.* **15**: 326–332.

PLASTERK, R. H. A., and H. G. A. M. VAN LUENEN, 1997 Transposons, pp. 97-116 in *C. elegans II*, edited by D. L. RIDDLE, T. BLUMENTHAL, B. J. MEYER and J. R. PRIESS. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

POLLAK, E., 1987 On the theory of partially inbreeding finite populations. I. partial selfing. *Genetics* **117:** 353–360.

REZSOHAZY, R., H. G. A. M. VAN LUENEN, R. M. DURBIN, and R. H. A. PLASTERK, 1997 Tc7, a Tc1-hitch hiking transposon in *Caenorhabditis elegans*. *Nucleic Acids Res.* **25**: 4048–4054.

RICCI, C., 1998 Anhydrobiotic capabilities of bdelloid rotifers. *Hydrobiologia* **388**: 321–326.

RICE, W. R., 1994 Degeneration of a nonrecombining chromosome. *Science* **263**: 230–232.

RIESEBERG, L. H., C. T. PHILBRICK, P. E. PACK, M. A. HANSON, and P. FRITSCH, 1993 Inbreeding depression in androdioecious populations of *Datisca glomerata* (Datiscaceae). *Am. J. Bot.* **80**: 757–762.

RISPE, C., and N. A. MORAN, 2000 Accumulation of deleterious mutations in endosymbionts: Muller's ratchet with two levels of selection. *Am. Nat.* **156**: 425–441.

RIZZON, C., G. MARAIS, M. GOUY and C. BIEMONT, 2002 Recombination rate and the distribution of transposable elements in the *Drosophila melanogaster* genome. *Genome Res.* **12**: 400–407.

RIZZON, C., E. MARTIN, G. MARAIS, L. DURET, L. SEGALAT *et al.*, 2003 Patterns of selection against transposons inferred from the distribution of Tc1, Tc3 and Tc5 insertions in the mut-7 line of the nematode *Caenorhabditis elegans*. *Genetics* **165**: 1127–1135.

ROCHELEAU, G., and S. LESSARD, 2000 Stability analysis of the partial selfing selection model. *J. Math. Biol.* **40**: 541–574.

ROZAS, J., J. C. SANCHEZ-DELBARRIO, X. MESSEGUER, and R. ROZAS, 2003 DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496–2497.

SÁNCHEZ-GARCIA, A., X. MASIDE and B. CHARLESWORTH, 2005 High rate of horizontal transfer of transposable elements in *Drosophila*. *Trends Genet*. **21**: 200–203.

SAWYER, S. A., D. E. DYKHUIZEN, R. F. DUBOSE, L. GREEN, T. MUTANGADURA-MHLANGA *et al*., 1987 Distribution and abundance of insertions sequences among natural isolates of *Escherichia coli*. *Genetics* **115**: 51–63.

SCHIERUP, M. H., and F. B. CHRISTIANSEN, 1996 Inbreeding depression and outbreeding depression in plants. *Heredity* **77**: 461–468.

SCHLENKE, T. A., and D. J. BEGUN, 2004 Strong selective sweep associated with a transposon insertion in *Drosophila simulans*. *Proc. Natl. Acad. Sci.* USA **101**: 1626–1631.

SCHMITT, J., and S. E. GAMBLE, 1990 The effect of distance from the parental site on offspring performance in *Impatiens capensis*: a test of the local adaptation hypothesis. *Evolution* **44**: 2022–2030.

SCHOEN, D. J., and A. H. D. BROWN, 1991 Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants. *Proc. Natl. Acad. Sci*. USA **88**: 4494–4497.

SCHOEN, D. J., M. O. JOHNSTON, A.-M. L'HEUREUX, and J. V. MARSOLAIS, 1997 Evolutionary history of the mating system in *Amsinckia* (Boraginaceae). *Evolution* **51**: 1090–1099.

SCHULENBURG, H., and S. MÜLLER, 2004 Natural variation in the response of *Caenorhabditis elegans* towards *Bacillus thuringiensis*. *Parasitology* **128**: 433–443.

SEIDEL, H. S., M. V. ROCKMAN, and L. KRUGLYAK, 2008 Widespread genetic incompatibility in *C. elegans* maintained by balancing selection. *Science*, in press.

SHELDAHL, L. A., D. M. WEINREICH, and D. M. RAND, 2003 Recombination, dominance and selection on amino acid polymorphism in the *Drosophila* genome: contrasting patterns on the X and fourth chromosomes. Genetics **165**: 1195–1208.

SHOOK, D.R., and T. E. JOHNSON, 1999 Quantitative trait loci affecting survival and fertility-related traits in *Caenorhabditis elegans* show genotype-environment interactions, pleiotropy and epistasis. *Genetics* **153**: 1233–1243.

SILVA, J. C., E. L. LORETO and J. B. CLARK, 2004 Factors that affect the horizontal transfer of transposable elements. *Curr. Issues Mol. Biol*. **6**: 57-71.

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.

SLATKIN, M., 1985 Genetic differentiation of transposable elements under mutation and unbiased gene conversion. *Genetics* **110**: 145–158.

SMITH, C. E., B. LLORENTE and L. S. SYMINGTON, 2007 Template switching during break-induced replication. *Nature* **447**: 102–105.

SNIEGOWSKI, P., and B. CHARLESWORTH, 1994 Transposable element numbers in cosmopolitan inversions from a natural population of *Drosophila melanogaster*. *Genetics* **137**: 815–827.

SÖDERBERG, R. J., and O. G. BERG, 2007 Mutational interference and the progression of Muller's ratchet when mutations have a broad range of deleterious effects. *Genetics* **177**: 971–986.

SOKAL, R. R., and F. J. ROHLF, 1995 *Biometry*. W. H. Freeman and Company: New York.

STEBBINS, G. L., 1957 Self fertilization and population variability in the higher plants. *Am. Nat.* **91**: 337–354.

STEIN, L. D., Z. BAO, D. BLASIAR, T. BLUMENTHAL, M. R. BRENT *et al.*, 2003 The genome sequence of *Caenorhabditis briggsae*: A platform for comparative genomics. *PLoS Biol.* **1**: e45.

STEINEMANN, M., and S. STEINEMANN, 1998 Enigma of Y chromosome degeneration: neo-Y and neo-X chromosomes of *Drosophila miranda* a model for sex chromosome evolution. *Genetica* **102/103**: 409–420.

STEINEMANN, S., and M. STEINEMANN, 2005 Y chromosomes: born to be destroyed. *BioEssays* **27**: 1076–1083.

STEPHAN, W., and Y. KIM, 2002 Recent application of diffusion theory to population genetics in *Modern Developments in Theoretical Population Genetics*, edited by M. SLATKIN and M. VEUILLE. Oxford University Press, Oxford, UK.

STEWART, A. D., and P. C. PHILLIPS, 2002 Selection and the maintenance of androdioecy in *Caenorhabditis elegans*. *Genetics* **160**: 975–982.

STEWART, M. K., N. L. CLARK, G. MERRIHEW, E. M. GALLOWAY, J. H. THOMAS, 2005 High genetic diversity in the chemoreceptor superfamily of *Caenorhabditis elegans*. *Genetics* **169**: 1985–1996.

SUH, D. S., E. H. CHOI, T. YAMAZAKI and K. HARADA, 1995 Studies on the transposition rates of mobile genetic elements in a natural population of *Drosophila melanogaster*. *Mol. Biol. Evol.* **12:** 748–758.

SULSTON, J., and J. HODGKIN, 1988 Methods. pp. 587-606 in *The Nematode Caenorhabditis elegans*, edited by W. B. WOOD. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

TAJIMA, F., 1989 Statistical method for testing the neutral mutation hypothesis. *Genetics* **123**: 585–595.

TAKASAKI, N., T. Y. AMAKI, M. HAMADA, L. PARK and N. OKADA, 1997 The salmon *Sma*I family of short interspersed repetitive elements (SINEs):

interspecific and intraspecific variation of the insertion of SINEs in the genomes of chum and pink salmon. *Genetics* **146**: 369–380.

TAKEBAYASHI, N., and P. P. MORRELL, 2001 Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. *Am. J. Bot.* **88**: 1143–1150.

TAM, S. M., M. CAUSSE, C. GARCHERY, H. BURCK, C. MHIRI *et al.*, 2007 The distribution of *copia*-type retrotransposons and the evolutionary history of tomato and related wild species. *J. Evol. Biol.* **20**: 1056–1072.

TEMPLETON, A. R., 1986 Coadaptation and outbreeding depression, pp. 33-59 in *Conservation Biology: The Science of Scarcity and Diversity*, edited by M. SOULÉ. Sinauer Associates, Sunderland, Mass.

TEOTÓNIO, H., D. MANOEL and P. C. PHILLIPS, 2006 Genetic variation for outcrossing among *Caenorhabditis elegans* isolates. *Evolution* **60**: 1300–1305.

THEODOROU, K., and D. COUVET, 2002 Inbreeding depression and heterosis in a structured population; influence of the mating system. *Genet. Res.* **80**: 107–116.

TRUYENS, S., M. M. ARBO, and J. S. SHORE, 2005 Phylogenetic relationships, chromosome and breeding system evolution in *Turnera* (Turneraceae): Inferences from its sequence data. *Am. J. Bot.* **92**: 1749–1758.

UYENOYAMA, M. K., K. E. HOLSINGER, and D. M. WALLER, 1993 Ecological and genetic factors directing the evolution of self-fertilization. *Oxf. Surv. Evol. Biol.* **9**: 327–381.

VAN LUENEN, H. G. A. M., S. D. COLLOMS and R. H. A. PLASTERK, 1994 The mechanism of transposition of Tc3 in *C. elegans*. *Cell* **79**: 293–301.

VAN VOORHIES, W. A., J. FUCHS and S. THOMAS, 2005 The longevity of *Caenorhabditis elegans* in soil. *Biol. Lett.* **1**: 247–249.

VASSILIEVA, L. L., and M. LYNCH, 1999 The rate of spontaneous mutation for life-history traits in *Caenorhabditis elegans*. *Genetics* **151**: 119–129.

VIEIRA, C., and C. BIÉMONT, 1997 Transposition rate of the 412 retrotransposable element is independent of copy number in natural populations of *Drosophila simulans*. *Mol. Biol. Evol.* **14**: 185–188.

VINEY, M. E., M. P. GARDNER, and J. A. JACKSON, 2003 Variation in *Caenorhabditis elegans* dauer larva formation. *Dev. Growth Differ.* **45**: 389–396.

VOGLER, D. W., K. FILMORE, and A. G. STEPHENSON, 1999 Inbreeding depression in *Campanula rapunculoides L*. I. A comparison of inbreeding depression in plants derived from strong and weak self-incompatibility phenotypes. *J. Evol. Biol.* **12**: 483–494.

VOS, J. C., and R. H. A. PLASTERK, 1994 Tc1 transposase of *Caenorhabditis elegans* is an endonuclease with a bipartite DNA binding domain. *EMBO J.* **13**: 6125–6132.

WAKELEY, J., 2003 Polymorphism and divergence for island-model species. *Genetics* **163**: 411–420.

WALLER, D. M., 1993 The statics and dynamics of mating system evolution. Pp. 97-117 *in* N. W. Thornhill, ed. The Natural History of Inbreeding and Outbreeding: Theoretical and Empirical Perspectives. University of Chicago Press, Chicago, IL.

WANG, J., W. G. HILL, D. CHARLESWORTH and B. CHARLESWORTH, 1999 Dynamics of inbreeding depression due to deleterious mutations in small populations: mutation parameters and inbreeding rate. *Genet. Res.* **74**: 165–178.

WASER, N. M., 1993 Sex, mating systems, inbreeding, and outbreeding, pp. 1-13 in *The natural history of inbreeding and outbreeding: theoretical and empirical perspectives*, edited by N. W. THORNHILL. University of Chicago Press, Chicago, IL.

WASER, N. M., and M. V. PRICE, 1994 Crossing-distance effects in *Delphinium nelsonii*: outbreeding and inbreeding depression in progeny fitness. *Evolution* **48**: 842–852.

WEEKS, S. C., 2004 Levels of inbreeding depression over seven generations of selfing in the androdioecious clam shrimp, *Eulimnadia texana*. *J. Evol. Biol.* **17**: 475–484.

WEEKS, S. C., V. MARCUS, and B. R. CROSSER, 1999 Inbreeding depression in a self-compatible, androdioecious crustacean, *Eulimnadia texana*. *Evolution* **53**: 472–483.

WEEKS, S. C., C. BENVENUTO and S. K. REED, 2006a When males and hermaphrodites coexist: a review of androdioecy in animals. *Integr. Comp. Biol.* **46**: 449–464.

WEEKS, S. C., T. F. SANDERSON, S. K. REED, M. ZOFKOVA, B. KNOTT, U. BALARAMAN, G. PEREIRA, D. M. SENYO, and W. R. HOEH, 2006b Ancient androdioecy in the freshwater crustacean *Eulimnadia*. *Proc. R. Soc. Lond.* B **273**: 725–734.

WEIR, B. S., and W. G. HILL, 1980 Effect of mating structure on variation in linkage disequilibrium. *Genetics* **95**: 477–488.

WHITE, M. J. D., 1973 *Animal Cytology and Evolution*. Cambridge University Press, Cambridge, UK.

WHITLOCK, M. C., 2002 Selection, load and inbreeding depression in a large metapopulation. *Genetics* **160**: 1191–1202.

WILLIS, J. H., 1999 The role of genes of large effect on inbreeding depression in *Mimulus guttatus*. *Evolution* **53**: 1678–1691.

WILSON, D. S., and M. TURELLI, 1986 Stable underdominance and the evolutionary invasion of empty niches. *Am. Nat.* **127**: 835–850.

WONG, A., P. BOUTIS, and S. HEKIMI, 1995 Mutations in the clk-1 gene of *Caenorhabditis elegans* affect developmental and behavioral timing. *Genetics* **139**: 1247–1259.

WRIGHT, S., 1977 *Evolution and the Genetics of Populations. Vol. 3. Experimental Results and Evolutionary Deductions*. University of Chicago Press, Chicago, IL.

WRIGHT, S. I., and D. J. FINNEGAN, 2001 Genome Evolution: Sex and the transposable element. *Curr. Biol.* **11**: R296–R299.

WRIGHT, S. I., and D. J. SCHOEN, 1999 Transposon dynamics and the breeding system. *Genetica* **107**: 139–148.

WRIGHT, S. I., Q. HIEN LE, D. J. SCHOEN and T. E. BUREAU, 2001 Population dynamics of an *Ac*-like transposable element in self- and cross-polinating *Arabidopsis*. *Genetics* **158**: 1279–1288.

WRIGHT, S. I., N. AGRAWAL and T. E. BUREAU, 2003 Effects of recombination rate and gene density on transposable element distributions in *Arabidopsis thaliana*. *Genome Res*. **13**: 1897–1903.

ZAMPICININI, G., A. BLINOV, P. CERVELLA, V. GURYEV and G. SELLA, 2004 Insertional polymorphism of a non-LTR mobile element (NLRCth1) in European populations of *Chironomus riparius* (Diptera, Chironomidae) as detected by transposon insertion display. *Genome* **47**: 1154–1163.

ZEYL, C., and G. BELL, 1996 Symbiotic DNA in eukaryotic genomes. *Trends Ecol. Evol*. **11**: 10–15.

Copyright © 2006 by the Genetics Society of America
DOI: 10.1534/genetics.106.060434

The Fate of Transposable Elements in Asexual Populations

Elie S. Dolgin[1] and Brian Charlesworth

Institute of Evolutionary Biology, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom

Manuscript received May 9, 2006
Accepted for publication July 31, 2006

ABSTRACT

Sexual reproduction and recombination are important for maintaining a stable copy number of transposable elements (TEs). In sexual populations, elements can be contained by purifying selection against host carriers with higher element copy numbers; however, in the absence of sex and recombination, asexual populations could be driven to extinction by an unchecked proliferation of TEs. Here we provide a theoretical framework for analyzing TE dynamics under asexual reproduction. Analytic results show that, in an infinite asexual population, an equilibrium in copy number is achieved if no element excision is possible, but that all TEs are eliminated if there is some excision. In a finite population, computer simulations demonstrate that small populations are driven to extinction by a Muller's ratchet-like process of element accumulation, but that large populations can be cured of vertically transmitted TEs, even with excision rates well below transposition rates. These results may have important consequences for newly arisen asexual lineages and may account for the lack of deleterious retrotransposons in the putatively ancient asexual bdelloid rotifers.

TRANSPOSABLE elements (TEs) are mobile DNA sequences that are abundant in the genomes of nearly all living organisms, including bacteria, protists, fungi, plants, and animals (Craig *et al.* 2002). Although there are cases in which mobile genetic elements may be coopted to serve host regulatory or structural functions (Kidwell and Lisch 2001), like most other mutator mechanisms, TEs are known to reduce the fitness of their host organism. However, unlike other classes of mutation, TEs are capable of autonomous self-replication. When elements transpose, they replicate faster than their host genome, with rates of transposition above rates of spontaneous deletion (Charlesworth and Langley 1989; Nuzhdin and Mackay 1995; Maside *et al.* 2000). This permits TE persistence despite their deleterious effects, as postulated by the "selfish DNA" hypothesis (Doolittle and Sapienza 1980; Orgel and Crick 1980).

Experimental and theoretical studies suggest that the ability of TEs to propagate can be balanced by natural selection against individuals with high element copy number (Charlesworth and Charlesworth 1983; Kaplan and Brookfield 1983). Three main sources of deleterious effects on fitness of segregating TEs have been postulated (see review by Nuzhdin 1999): insertions disrupting gene function (Finnegan 1992), chromosomal rearrangements generated by ectopic exchange (Montgomery *et al.* 1987), and a selective cost of transposition itself (Brookfield 1991). Host regulatory mechanisms may play a role as well, and there may be rare beneficial insertions, but it is likely that nearly all TE insertions are deleterious. There is substantial evidence from natural populations and experimental studies for such deleterious effects (Charlesworth *et al.* 1994).

Coevolution between TEs and their hosts depends critically on the presence of sexual reproduction. Outcrossing provides a means for TEs to spread to all individuals in a population. Despite being deleterious, TEs can persist in a sexual population while inflicting a severe fitness penalty on their hosts because they can increase in number faster than the host genome (Hickey 1982). This would not be true in an asexual population. In the absence of horizontal transmission, there is no between-lineage transmission, and TEs cannot initially spread in an asexual population (Hickey 1982). However, asexual lineages generally arise from sexual progenitors whose genomes are riddled with TEs. Upon the abandonment of sex, elements are coupled entirely to their hosts. The proliferation of TEs would be detrimental to the host to which they are confined. In the long run, TEs would be expected to become inactive and domesticated, as lineages with inactive elements should have a selective advantage and outcompete other lineages (Doolittle *et al.* 1984). However, theoretical studies show that TEs in outcrossing populations will generally evolve maximum transposition rates (Charlesworth and Langley 1986). Asexual lineages that arise from sexual populations will contain actively transposing TEs, selected to multiply without considering the fitness of the ancestral sexual host. By doing so, they would seal the fates of their asexual hosts. Accordingly, a long-term

[1]*Corresponding author:* Institute of Evolutionary Biology, University of Edinburgh, Ashworth Laboratories, King's Bldgs., W. Mains Rd., Edinburgh EH9 3JT, United Kingdom. E-mail: elie.dolgin@ed.ac.uk

advantage of sex may result from the early extinction of asexual lineages, due to the unchecked proliferation of TEs upon the abandonment of sex (Arkhipova and Meselson 2005a).

Despite the fact that sex facilitates the spread of TEs (Hickey 1982), it may thus also be necessary to contain their proliferation. In this article, we examine this idea by modeling the population dynamics of transposable elements under asexual reproduction. While the abandonment of sex in small populations will probably lead to element buildup and eventual host extinction in small populations, large populations with some level of element excision or deletion may be able to eliminate vertically transmitted deleterious TEs, providing a potential benefit of asexuality.

METHODS

Assumptions of the models: We consider a diploid asexual lineage arising from an ancestral sexual population that contains deleterious TEs. With offspring genotypes identical to those of their parents, we assume that a clonal lineage originates from a single ancestor, such that initially all individuals have a TE copy number x. Elements are transmitted only vertically from parent to offspring. The copy number may increase due to transposition, at a rate u per element, or decrease due to excision at a rate v per element. We first consider analytic models of TE dynamics in infinite asexual populations. We then employ a stochastic simulation method to investigate the effects of finite population sizes.

Infinite populations: To examine TE dynamics with infinite population size, we adapt the exact recurrence relations of Kimura and Maruyama (1966; Equation 3.1), originally devised to calculate the mutation load in asexual populations. We assume an infinite number of insertion sites and Poisson distributions of the numbers of transposition and excision events as an approximation for binomial distributions. We calculate the transposition load in the presence and absence of excision, without specifying the relation between fitness and copy numbers. We then consider specific fitness functions and determine the mean and variance element of copy number using iterations and approximations based on the method of Charlesworth (1990).

Finite populations: Computer simulations were used to examine the effects of excision in finite asexual populations. The population size N is constant, with offspring genomes identical to those of their parents except for transposition and excision events. The genome is made up of two diploid chromosomes, with each chromosome able to carry 200 elements, although since the population is asexual and nonrecombining, this is equivalent to any number of chromosomes of any ploidy level. A simulation run is initiated with a single individual carrying x elements, which founds the clonal population, such that all N individuals are initially identical. TE dynamics are then monitored over many generations, where each generation reproduces asexually, followed by transposition and excision. Reproduction involves randomly sampling individuals to produce a new offspring population of size N, where the probability of an individual being selected as a parent is proportional to its fitness. As in the case of the infinite population size model, the numbers of transposition and excision events were drawn from Poisson distributions, with constant probabilities per element of u and v, respectively. New insertions were placed randomly in the genome at unoccupied sites, and the exact locations of elements were maintained between generations unless they had been excised.

Simulations were run until all elements had been lost from the population or until the mean TE copy number had accumulated substantially (>150). Above this number, a runaway process of element buildup is observed (results not shown). No stable equilibrium with an intermediate TE copy number was ever observed, consistent with results from an asexual model of TE sequence evolution for simulations in which TE inactivation through mutation was not included (Docking *et al.* 2006). We calculated the proportion of simulation runs where TEs were eliminated, to examine the conditions under which asexual populations are expected to cure themselves of deleterious TEs or be driven to early extinction. We also monitored the variance in copy number and the minimum TE copy number in the population. C++ files of this simulation program are available upon request. The random numbers were implemented using the Mersenne Twister pseudorandom number generator (Matsumoto and Nishimura 1998), adapted for C++ by J. Bedaux (http://www.bedaux.net/mtrand/).

Selection: Fitness was assumed to be a decreasing function of TE abundance, as would be expected in the ancestral population. In sexual populations, a stable equilibrium copy number occurs only when there are synergistic fitness interactions between elements (Charlesworth and Charlesworth 1983). Thus, the fitness of an individual with n elements was represented by an exponential quadratic, decreasing function of the copy number,

$$w_n = \exp(-an - \tfrac{1}{2}bn^2), \qquad (1)$$

where a and b are constant selection coefficients (Charlesworth 1990).

Initial copy number of clonal lineage: We set the initial copy number, x, to be equal to the equilibrium mean copy number of the ancestral sexual population. In outcrossing populations, with low frequencies of elements at each occupied site, copy number dynamics are described by

$$\Delta\bar{n} \approx V_n \frac{\partial \ln w_{\bar{n}}}{\partial \bar{n}} + \bar{n}(u - v), \qquad (2)$$

where $\bar{n}$ is the mean population copy number, V_n is the variance in copy number, and $w_{\bar{n}}$ is the fitness of an individual carrying the mean number of elements (Charlesworth 1985). Solving Equation 2 for $\Delta\bar{n} = 0$, assuming a Poisson distribution of elements, and using the quadratic fitness formula, the equilibrium mean element copy number of the ancestral sexual population is

$$\bar{n} \approx (u - a - v)/b. \quad (3)$$

The transposition load of a sexual population is measured as

$$L = -\ln \bar{w} = a\bar{n} + \tfrac{1}{2}b\bar{n}^2. \quad (4)$$

However, due to the synergism in fitness effects, there is a departure from independence among elements, resulting in the equilibrium load due to TEs being approximately half that of Equation 4 when $\bar{n}$ is given by Equation 3 (Kimura and Maruyama 1966; Crow 1970; Charlesworth 1990; Charlesworth and Barton 1996). We thus set the equilibrium load, $L_{\mathrm{eq}} = \frac{1}{2}L$, and solve for the equilibrium mean copy number under synergistic fitness, such that

$$n_{\mathrm{eq}} = \frac{\sqrt{2(a^2 + (u - v)^2)} - 2a}{2b}. \quad (5)$$

Model parameters: We set the parameters in our simulations by using estimates from studies of Drosophila populations. Average transposition rates per element per generation are $\sim 10^{-4}$, regardless of the class of element, with excision rates at least one order of magnitude smaller (Nuzhdin and Mackay 1995; Vieira and Biémont 1997; Pasyukova *et al.* 1998; Maside *et al.* 2000). We set $u = 10^{-4}$ and investigate excision rates of $v = 10^{-5}$ and $v = 10^{-6}$. The strength of selection on segregating elements in natural populations should be of the order 10^{-5}–10^{-4} per copy, if no insertions are completely neutral (Charlesworth *et al.* 1994). With synergism between elements, the strength of selection will depend on both a and b. We set $a = 10^{-5}$ and let b vary according to Equation 5 for a given n_{eq}. We set the initial copy number, x, of the founding clone to n_{eq} and assume that asexual lineages will experience the same parameters as the sexual ancestor. We use these values of transposition and selection to investigate the impact of population size on different excision rates.

RESULTS

Infinite populations without excision: We modify the exact recurrence relations of Kimura and Maruyama (1966; Equation 3.1) to analyze TE dynamics in an infinite asexual population without element excision. We write f_i for the frequency of individuals with i elements, whose fitness is w_i. We assume a Poisson distribution of transposition insertions with mean ui in a genome with i elements, where each transposition event leads to a new insertion. With no excision ($v = 0$), the frequency of individuals with i elements in the next generation is

$$f_i' = \frac{1}{\bar{w}} \sum_{j=0}^{i} f_j w_j \frac{(uj)^{i-j}}{(i-j)!} e^{-uj}, \quad (6)$$

where $\bar{w} = \sum f_i w_i$ is the mean fitness. Since there is no excision to reduce the TE copy number below that of the initial clone, the frequency of individuals with the lowest copy number is f_x. Following the derivation of Kimura and Maruyama (1966),

$$f_x' = \frac{f_x w_x}{\bar{w}} e^{-ux}, \quad (7)$$

and the lineage reaches an equilibrium with mean fitness

$$\bar{w} = w_x e^{-ux}. \quad (8)$$

Since individuals with the lowest element copy number have the highest fitness, all else being equal, a stable equilibrium is achieved with a transposition load,

$$L = ux. \quad (9)$$

This is equivalent to the mutational load of an asexual population (Kimura and Maruyama 1966), where the genomic mutation rate, U, is replaced by the product of the transposition rate and initial copy number. Thus, the load depends critically on the initial number of TEs in the founding clone. This result is quite general and independent of the fitness function.

The equilibrium mean and variance in element copy number can be calculated by iteration, given a specific fitness function, transposition rate, and initial copy number. Using Equation 1, a normal-distribution approximation can be used to estimate the equilibrium solutions. We adapt Equations 2–4 of Charlesworth (1990), originally developed to calculate the equilibrium number of mutations in an asexual population under mutation–selection balance, to determine the equilibrium properties of transposable elements under asexuality without excision. Using his method of assuming an approximately normal distribution of copy number among individuals in the population, at equilibrium we have

$$\hat{n} = \frac{aV_n}{u + bV_n(u - 1)} \quad (10)$$

and

$$2\exp(-x(u + a + \tfrac{1}{2}bx)) = \ln(1 + bV_n) - \frac{1}{1 + bV_n}(a^2 V_n - 2a\hat{n} - b\hat{n}^2), \quad (11)$$

where $\hat{n}$ is the mean asexual population TE copy number and V_n is the variance in copy number.

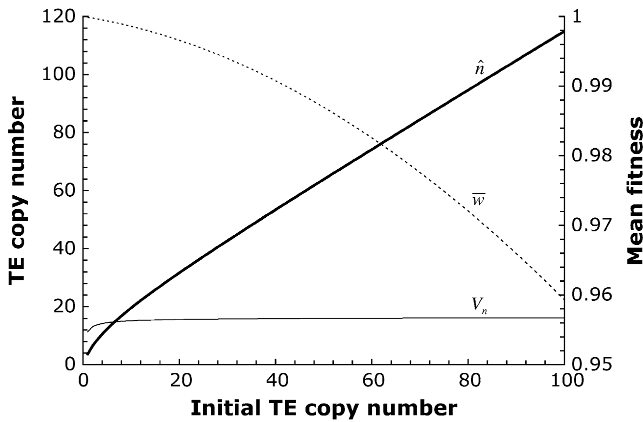


FIGURE 1.—Equilibrium mean properties of an asexual population in the absence of excision, at transposition–selection balance with a quadratic fitness function. The primary *y*-axis shows how the mean copy number (thick line) and variance in copy number (thin line) vary with the initial copy number at equilibrium. The dashed line uses the secondary *y*-axis to denote the mean fitness of the equilibrium population. The parameter values used are: rate of transposition per element $u = 10^{-4}$, selection coefficient $a = 10^{-5}$, and synergism coefficient $b = 6.1 \times 10^{-6}$.

These equations allow the mean and variance in TE copy number at equilibrium to be calculated by iteration (Charlesworth 1990). We set $u = 10^{-4}$, $a = 10^{-5}$, $n_{eq} = 10$ and evaluate b using Equation 5, giving $b = 6.1 \times 10^{-6}$. We then calculate the equilibrium solutions for an asexual population, with TE copy numbers in the initial clone from 1–100. Figure 1 shows the equilibrium properties of an infinite asexual population for the given set of parameters solved, using the normal approximation and confirmed by iteration of Equation 6. A stable equilibrium is always achieved, with a mean copy that depends nearly linearly on *x*. The mean fitness declines with larger *x*, while the variance in copy number is nearly constant.

Unlike sexual populations, where the logarithm of fitness must decline faster than linearly with TE copy number (Charlesworth and Charlesworth 1983), a stable equilibrium can be achieved in an asexual population with an arbitrary fitness function. We now consider selection in the absence of synergism ($b = 0$), as may be the case if these interactions are dependent on the presence of sexual reproduction and recombination (*e.g.*, with meiotic ectopic exchange; Langley *et al.* 1988). Without synergism, the mean and variance in TE copy number simplify to

$$\hat{n} = \frac{2x(a+u)}{a(2-u)} \approx \frac{a+u}{a}x$$
$$V_n = \frac{u}{a}\hat{n} \qquad (12)$$

These equations were confirmed by comparison with the exact equilibrium solutions obtained by iteration of Equation 6. It can be seen that the mean and variance scale linearly with the initial TE copy number, *x*. If the rate of transposition is much greater than the strength of selection against segregating TEs, the mean copy number and variance can be exceedingly high in the absence of synergism. For example, with the parameters used previously ($u = 10^{-4}$, $a = 10^{-5}$), $\hat{n} \approx 11x$ and $V_n \approx 10\hat{n}$. Thus, while a stable equilibrium is theoretically achieved in the absence of synergism between elements, such fitness interactions may be necessary to hold copy number at a biologically relevant equilibrium.

Infinite populations with excision: With element excision, the previous analysis no longer holds. The lowest copy number can be reduced below that of the initial clone, and the individuals with fewer TEs should have a selective advantage, leading to a decline in the mean number of elements in the population. Regardless of the TE count of the initial clone, in an infinite asexual population, it is eventually possible for all TEs to be excised, creating a class of individuals without any elements. Unlike the situation with recurrent mutations, once such a zero class has been generated, it is immune to further transposition, assuming that there is no horizontal transfer. Due to its fitness advantage, this class will spread to fixation, effectively curing the population of deleterious TEs. This was confirmed by modifying Equation 6 to include excision and solving by iteration. The mean and variance in TE copy number always decline to zero irrespective of the rates of transposition and excision, the fitness function, or the initial copy number.

Parameter scaling of finite populations: To further investigate whether asexual populations achieve the equilibrium expectations without excision and whether populations with excision can eliminate all their deleterious TEs, computer simulations were used to examine the effects of finite population sizes. Running simulations using the biologically realistic parameters derived from Drosophila populations is very time-consuming, because of the large population sizes and small transposition rates, excision rates, and selection coefficients. According to population genetics theory, if evolutionary forces are weak, so that diffusion approximations can be used, the properties of the system are determined by the values of the products of the deterministic forces and the effective population size (Ewens 1979, Chap. 4). If these values are maintained constant, we should expect the same evolutionary outcomes.

We confirmed this expectation by comparing simulations with the parameters from Drosophila and an adjusted version of these parameters for $N = 10^4$, such that $Nu = 1$, with the other ratios kept constant as well. To achieve this, we scale up *u*, *v*, *a*, and *b* one and two orders of magnitude while scaling down *N* and the numbers of generations by the same amount. The accuracy of this parameter adjustment is demonstrated in Figure 2. By reducing the population size and the time of the simulation accordingly, this simplification seems to be reasonably accurate, although perhaps somewhat

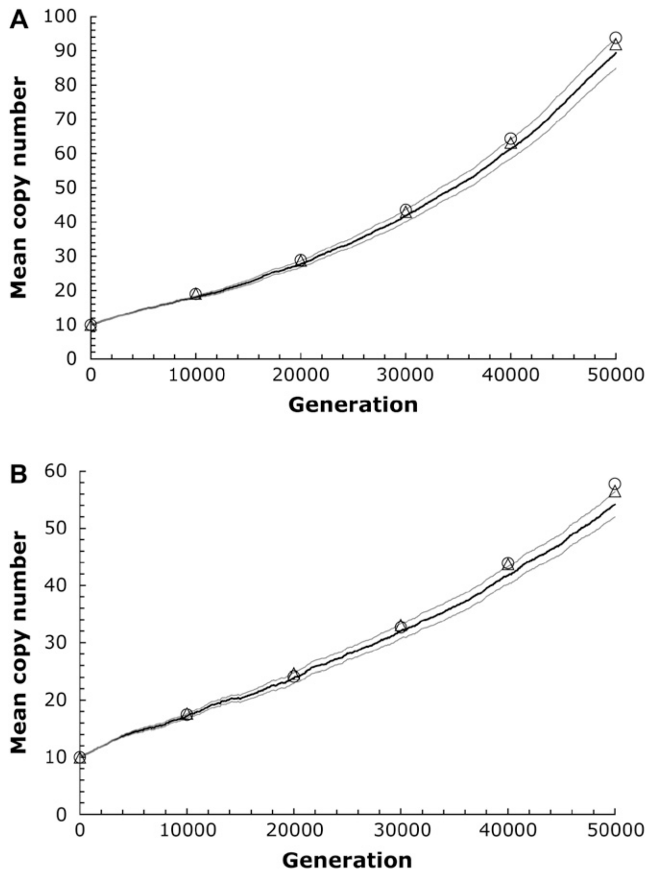


FIGURE 2.—Demonstrating the accuracy of scaling parameters. The thick line shows the mean copy number from computer simulations run over 50,000 generations for a population size $N = 10^4$, with initial copy number $x = 10$, rate of transposition per element $u = 10^{-4}$, selection coefficient $a = 10^{-5}$, and synergism coefficient b. The equilibrium is solved for a given rate of excision per element according to Equation 5. (A) An excision rate two orders of magnitude lower than transposition ($v = 10^{-6}$; $b = 6.0 \times 10^{-6}$; number of simulation runs, $r = 25$). (B) An excision rate one order of magnitude lower than transposition ($v = 10^{-5}$; $b = 5.4 \times 10^{-6}$; $r = 30$). The thin lines denote the standard error of the mean between runs of the simulation. The triangles show the mean copy number from simulations scaled 10-fold ($r = 100$). The circles show the mean copy number from simulations scaled 100-fold ($r = 1000$). The generation time is also scaled upward to compare with the unadjusted simulation runs.

overestimating the mean copy number, especially with higher excision rates. However, this allows us to perform a greater number of simulation runs in a fraction of the computation time and should be a conservative correction, since it potentially biases our results against finding situations when complete TE elimination is expected. We continued by using 100-fold greater transposition, excision, and selection values than those estimated from natural populations. This implies that the population sizes and times used in our simulations need to be scaled upward two orders of magnitude to reflect biologically relevant values. The results presented below show the uncorrected parameter values used in the computer simulations. However, since the magnitude of the population size is so critical to our results, we also include the scaled population sizes in parentheses.

Finite populations without excision: Unlike the case of infinite populations without excision, in simulations with $v = 0$, a large but finite population size, $N = 10^5$ (equivalent to 10^7 when rescaled), and using a scaled set of the parameters presented in Figure 1 ($u = 10^{-2}$, $a = 10^{-3}$, and $b = 6.1 \times 10^{-4}$), no equilibrium was ever observed for moderately low initial copy numbers, $x = 5$ and $x = 10$. Elements always accumulated substantially within several thousand generations (up to millions of generations, when rescaled), far above the infinite population equilibrium values of $\hat{n} = 12.3$ and $\hat{n} = 19.5$, respectively. This process of element accumulation is accelerated in smaller populations, with larger initial copy numbers, and with weaker selection. This is consistent with the results for TE sequence evolution in simulations with no excision, where no stable equilibrium in copy number was ever reached (DOCKING *et al.* 2006). This suggests that asexual populations with no excision of TEs are likely to accumulate large numbers of copies within a time frame of millions of generations, even if the effective population size is in the millions. Of course, with a sufficiently large population size, the population will spend a very long time near the equilibrium derived in the previous section, so that it will appear as though no accumulation is happening, but ultimately the population is expected to experience an unbounded proliferation of TEs.

Finite populations with excision: As the results of infinite populations with element excision showed, deleterious TEs can be completely purged from asexual populations when $v > 0$. The simulation results illustrate the interaction of population size and excision rate on the probability of TE elimination (Figure 3). In contrast to the expectation of complete element elimination in an infinite asexual population, TEs tend to accumulate substantially in small populations. Even with element excision, deleterious TEs accumulate in small populations well above the copy number expected at equilibrium in the absence of excision (see Figure 1). On the other hand, in large populations, the efficacy of selection and excision is enhanced. Consistent with the findings for the infinite population model, a sufficiently large asexual lineage can effectively purge itself of vertically transmitting deleterious TEs upon the abandonment of sex.

Two main factors contribute to the proliferation of TEs in small populations. As with sexual populations, this is partly due to the reduction in the variance in copy number between individuals in a finite population, thereby limiting the power of natural selection (BROOKFIELD and BADGE 1997). While we present results only on the final state of the populations, we also monitored the variance in copy number between generations over the course of the simulations. In comparisons

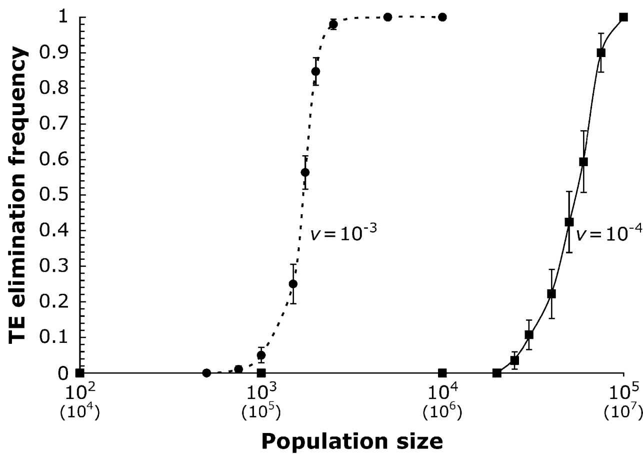


Figure 3.—TE elimination frequency and population size. The initial copy number was $x = 10$, $u = 10^{-2}$, and $a = 10^{-3}$; the synergism coefficient, b, was calculated for a given rate of excision per element using Equation 5. The solid line represents excision two orders of magnitude below transposition ($v = 10^{-4}$; $b = 6.0 \times 10^{-4}$). The dashed line represents excision one order of magnitude below transposition ($v = 10^{-3}$; $b = 5.4 \times 10^{-4}$). With $v = 10^{-4}$, all points with a TE elimination frequency of 0 or 1 are based on at least 10 simulation runs, and all points with an intermediate frequency include at least 30 simulation runs. With $v = 10^{-3}$, all points are based on at least 60 simulation runs. Error bars denote one standard error. To reflect plausible parameters for natural populations, all transposition and selection terms must be scaled down 100-fold. Scaled population sizes are shown in parentheses below the x-axis.

among populations that accumulate TEs, smaller populations indeed have a lower variance (results not shown). In addition, random genetic drift may lead to the loss of the class of individuals with the lowest copy number. In the absence of excision, the loss of this class would be irreversible, and TEs would continue to build up, leading to a decline in the mean fitness of the population. This process would be similar to that described for mutations, known as Muller's ratchet (Muller 1964; Stephan and Kim 2002), except that the genomic rate of transposition would increase with the copy number, leading to accelerated rates of fitness deterioration. With element excision, the process is reversible, but the rate at which the least-loaded class is lost may greatly exceed the rate of excision to generate a new best class. We kept track of the minimum TE copy number in the population during our simulations and observed the copy number of the least-loaded class to increase gradually above x as a population accumulated TEs (results not shown). Both of these factors probably contributed to the increased rate of TE proliferation in smaller populations.

In addition to the rate of excision relative to transposition, the ability of an asexual population to cure itself of deleterious TEs depends critically on the amount of synergism between elements, the magnitude of both transposition and excision, and the initial copy number to be eliminated. Synergism affects the ability of selection to limit TE proliferation in asexual populations. Higher rates of transposition result in faster TE proliferation, even if the ratio of transposition to excision is kept constant. The total number of initial TEs reduces the population mean fitness and affects the total genomewide rate of transposition. All of these parameters influence the ability of selection and excision to limit copy number.

To explore these factors, we looked at the effect of varying the degree of synergism, b, and changing the magnitude of transposition and excision, while maintaining the ratio of u:v constant at 10. None of these parameters are strictly independent, because they influence the expected equilibrium distribution of elements in the ancestral sexual population. We varied the initial copy number, x, from 10 to 50 and used Equation 5 to solve for the corresponding equilibrium value of b and the equilibrium values of u and v. In both cases, for the parameters that were held constant, we used the values employed previously. Figure 4 shows the effects of synergism, magnitude of transposition–excision, and initial copy number on the probability of TE elimination at two large population sizes.

All else being equal, increased synergism and lower transposition–excision rates reduce the population size necessary for TE elimination. Increased synergism strengthens selection against large copy numbers, and lower rates of transposition–excision slow the rate of TE accumulation, thereby improving the efficacy of selection. With strong synergism or low rates of transposition–excision, TEs are quite effectively removed from large populations regardless of the initial copy number. In contrast, with low initial copy numbers, TEs are likely to be eliminated, even with weak synergism or strong transposition–excision. Using parameters where intermediate frequencies of TE elimination are observed, a fine balance exists between selection and genetic drift. Slight differences in any of the parameters or the population size can shift this balance and greatly affect whether an asexual population increasingly accumulates or cures itself of deleterious TEs. Weaker synergism or higher transposition–excision rates also correspond to larger equilibrium mean copy numbers in an ancestral sexual population. Thus, the expected initial copy number will be greater, further hindering the ability of an asexual lineage to purge TEs. However, our approach may be viewed as conservative, since for all analyses other than that in Figure 4 we initiated the asexual population with the TE copy number, x, set to the sexual equilibrium. Given the variance in copy number in the ancestral population, however, an asexual lineage may arise from lower copy number individuals, thereby reducing x and increasing the probability of TE elimination. The asexual populations that survive are more likely to have arisen from low copy number founding populations.

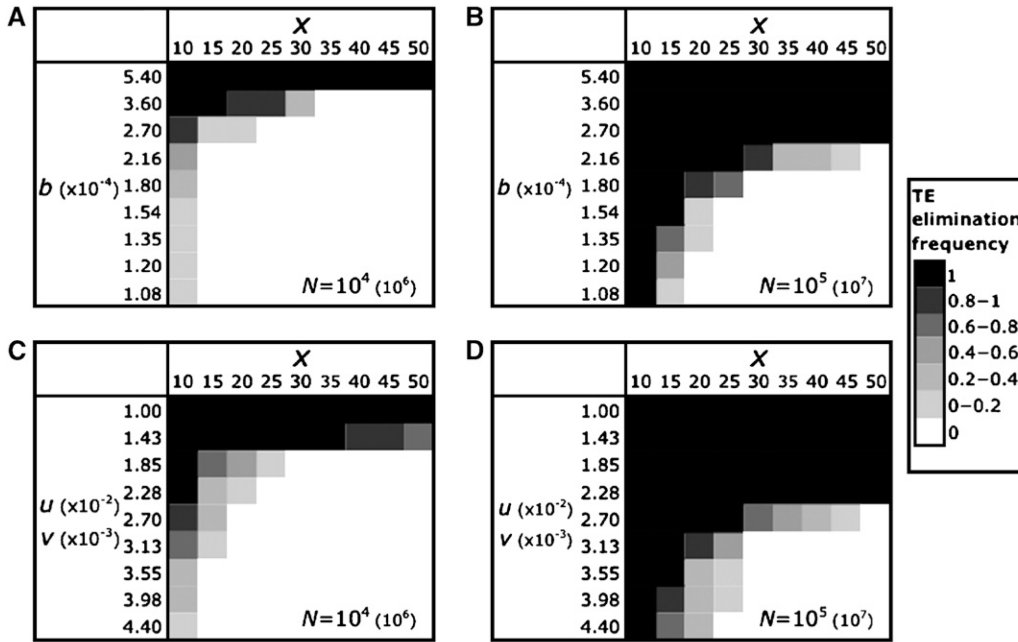


FIGURE 4.—The effect of varying the degree of synergism, magnitude of transposition and excision, and initial copy number on the frequency of TE elimination for two large population sizes. (A and B) The effect of changing b. (C and D) The effect of changing the magnitude of u and v while keeping the ratio of u:v constant at 10, while allowing x to vary. The ancestral sexual equilibrium solutions according to Equation 5 are located along the diagonal. The parameters that do not change are $u = 10^{-2}$, $v = 10^{-3}$, $a = 10^{-3}$, and $b = 5.4 \times 10^{-4}$. All points are based on at least 10 simulation runs. To reflect plausible parameters for natural populations, all transposition and selection terms must be scaled down 100-fold. Scaled population sizes are shown in parentheses next to the values used in the simulations.

Our model has so far supposed that all TE insertions are deleterious, but TEs may vary in their selective effects. In the absence of meiotic ectopic exchange, TE insertions in intergenic regions of the genome may not affect fitness if these regions have little or no functional role. We consider a model with two classes of insertion sites: those in which TEs affect fitness according to Equation 1 and those in which TEs are selectively neutral (CHARLESWORTH 1991). We assume that new TE insertions are neutral with a constant probability p and are deleterious with a probability $1 - p$. Figure 5 shows the effect of varying p on the TE elimination frequency. With a high proportion of neutral sites, asexual populations are far less likely to be able to eliminate TEs, since elements tend to build up at neutral sites and continue to transpose. This effect is magnified if TEs inherited from the ancestral sexual population are also present at neutral sites, as would be the case if TEs in this population were primarily contained by ectopic exchange, and hence found in sites where insertions have little or no direct fitness effects in the absence of recombination (MONTGOMERY *et al.* 1987; CHARLESWORTH *et al.* 1994).

DISCUSSION

Whether an asexual population can cure itself of deleterious TEs depends critically on the ability of selection and excision to limit the propagation of elements. The probability of a population achieving a TE-free state is influenced by the parameters controlling TE dynamics, including the scale of transposition and excision, and the degree of synergism between elements. However, all else being equal, the major factor affecting TE elimination is the population size of the asexual lineage. Since

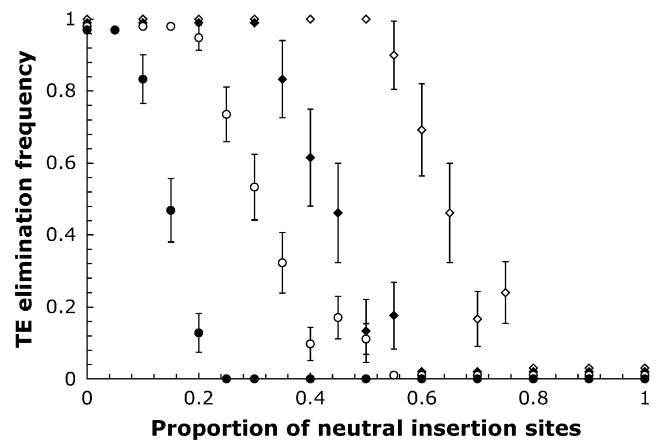


FIGURE 5.—TE elimination frequencies incorporating neutral sites. The open symbols show when all x sites containing the initial TEs in the founding clone are under selection; the solid symbols show when these are neutral. The circles show simulations done with $N = 10^4$, diamonds show simulations done with $N = 10^5$, and the remaining parameters are as before: $x = 10$, $u = 10^{-2}$, $v = 10^{-3}$, $a = 10^{-3}$, and $b = 5.4 \times 10^{-4}$. All symbols are based on at least 30 simulation runs for $N = 10^4$ and at least 10 simulation runs for $N = 10^5$. Error bars denote one standard error. To reflect plausible parameters for natural populations, all transposition and selection terms must be scaled down 100-fold, and the population size must be scaled up 100-fold.

the strength of selection on segregating elements is expected to be relatively low, selection will be ineffective in removing TEs unless the population size is large, and TEs will continue to accumulate at an ever-increasing rate.

Our results are in accordance with the hypothesis that deleterious TEs can drive the extinction of asexuals (Arkhipova and Meselson 2005a). However, we also provide a means by which an asexual lineage may be able to rid itself of the transposition load that it inherits upon the abandonment of sex. This possibility would provide a long-term benefit to asexual rather than sexual reproduction. Small populations may suffer from a reduced variance in TE copy number and a random loss of the class of individuals with low TE copy numbers. In large populations, however, selection is more effective and the Muller's ratchet-like process of element accumulation can be reversed because of excision. Even rates of TE excision well below that of transposition can improve overall fitness in large populations. Excision events are doubly advantageous, by both raising fitness and reducing the genomewide transposition rate because of reduced copy number. This allows mean fitness to improve faster than the stochastic loss of the least-loaded class, leading to the eventual elimination of all TEs. Once a class of individuals free of all TEs has been achieved, they are immune to further transposition and can spread through the population, arresting the process of TE-induced fitness deterioration. But retrotransposons notoriously show little or no excision (Craig *et al.* 2002), which poses a problem for this model that we consider below.

Most theoretical models of transposable elements are based on the biology of sexual eukaryotes, as selfish DNA can invade a population only if there is sexual exchange between individuals (Hickey 1982). The widespread presence of TEs in bacteria has prompted some theoretical analyses of TE dynamics in asexual populations. These models have shown that TE maintenance can be explained by frequent horizontal transmission (Sawyer *et al.* 1987; Hartl and Sawyer 1988), by potential positive impacts of TEs, including beneficial TE insertions causing the fixation of other neutral or slightly deleterious TEs in the genome by hitchhiking (Martiel and Blot 2002), or by selective advantages of TE insertions in fluctuating environments with environment-dependent selection (Edwards and Brookfield 2003). Basten and Moody (1991) obtained analytic expressions for equilibrium distributions of TEs in prokaryotic populations, using branching process theory and a number of selection models. In the absence of beneficial insertions, maintenance of TEs requires horizontal transfer, and the net rate of element change due to transposition and excision must be a decreasing function of TE copy number.

Here we focus exclusively on vertically transmitted deleterious TEs in eukaryotic asexual populations. This may be more representative of retroelements, for which horizontal transfer is extremely rare (Jordan *et al.* 1999; Malik *et al.* 1999), than of DNA-based transposons, for which there is frequent evidence for horizontal transfer (Silva *et al.* 2004). Previous authors have proposed that, with negligible transmission rates between individuals, TEs should either accumulate in asexual populations to the point of lineage extinction or be completely eliminated (Arkhipova and Meselson 2000, 2005a; Wright and Finnegan 2001; Nuzhdin and Petrov 2003). However, it remained unclear how asexual lineages could escape long-term degeneration and completely eliminate all the TEs in the genome. TEs can decay after accumulating mutations, but it seems likely that they will continue to multiply upon the abandonment of sex, given that the rate of inactivating mutations in TEs is probably similar to the rate of transposition (Nuzhdin *et al.* 1998). Docking *et al.* (2006) carried out simulations of TE sequence evolution after the abandonment of sex, including the effects of inactivating mutations, They found that, while all elements were eventually lost from asexual populations, it took many thousands of generations for selection on TE sequences to be relaxed, even with higher mutation and transposition rates than expected in nature. This suggests that new asexual lineages should indeed harbor many active TEs, perhaps for millions of generations assuming a scaling of parameters similar to that done in our model.

Nuzhdin and Petrov (2003) suggested that suppressor alleles in a lineage abandoning sex could result in the instantaneous inactivation of TEs. Indeed, our results for infinite populations imply that a lineage with a *trans*-acting repressor (located in either the host or the TE genomes) that reduces the rate of transposition, u, for all elements would have a fitness advantage, by reducing the load due to TEs, $L = ux$ (Equation 9). Nontransposing elements would eventually decay under mutation pressure. Our results provide, however, a simpler explanation for complete TE elimination, based on straightforward population genetic processes and biologically plausible parameters and requiring only large population sizes to enhance the efficacy of selection and excision.

The model presented here generally assumes that TEs are uniformly deleterious on a neutral background. The potential for TE elimination may be confounded by the presence of beneficial mutations, whether derived from TEs or otherwise, which may fix linked deleterious TEs by selective sweeps (Charlesworth *et al.* 1992). However, this requires selection coefficients for beneficial mutations to be substantially greater than those for deleterious TEs (Johnson and Barton 2002; Bachtrog and Gordo 2004). If excision occurs, this would only reduce the rate of TE elimination, by fixing a clonal lineage with a new initial copy number and lower mean fitness. A further limitation on TE elimination arises if a fraction of sites allow neutral TE insertions

(Charlesworth 1991). The continual insertion of TEs into such sites causes a buildup of elements and reduces the ability of selection to contain them at other sites in the genome.

Our model has investigated the eventual fate of an asexual lineage, either complete TE elimination or runaway element accumulation, but has not generally considered the time frame involved. In our simulations, the time to reach either final state generally ranges from several hundred generations to tens of thousands of generations, depending on the parameters involved (results not shown). However, the number of generations generally scales with the population size (see Figure 2). This implies that, with the parameter adjustment employed in our model, the time required to drive an asexual lineage to the point of extinction could be as high as millions of generations for larger populations. Similarly, complete TE elimination could take a very long time, especially with high initial copy numbers or weaker selection and excision. Thus, asexual lineages could persist with active deleterious TEs for a substantial period of time before arriving at either final state; however, on a long-term evolutionary scale (*i.e.*, tens of millions of years), it may be the case that only asexual populations that have purged all their deleterious TEs will persist (Arkhipova and Meselson 2005a).

Asexual lineages generally compose only a single "species" and rarely embrace a taxonomic group of higher rank. While asexual lineages that arise from sexual populations may prosper in the short term, they almost invariably suffer early extinction (see review by Judson and Normark 1996). Our results have important implications for the existence of ancient asexual taxa: those eukaryotic taxa that have evolved and reproduced for a long stretch of evolutionary time without sex or recombination and have been labeled as "evolutionary scandals" (Maynard Smith 1986). While many claims of ancient asexuality have been refuted upon further investigation, the most compelling evidence for ancient asexuality exists for the bdelloid rotifers, which are not known to have males, hermaphrodites, or meiosis (Arkhipova and Meselson 2005a). Strikingly, bdelloid rotifers appear to lack deleterious vertically transmitted retrotransposons. Unlike all sexually reproducing eukaryotic species that have been examined, including the monogonont rotifers, reverse transcriptases of two superfamilies of retrotransposons, gypsy-like and LINE-like retrotransposons, were not detectable in any of five bdelloid species (Arkhipova and Meselson 2000). This contrasts sharply with the patchy distribution of various DNA transposons found in bdelloid rotifers (Arkhipova and Meselson 2000, 2005b), in agreement with their often horizontal mode of transmission, and the presence of a domesticated retrotransposon (Arkhipova *et al.* 2003).

Presumably, the common sexual ancestor of the bdelloid and monogont rotifers harbored active deleterious TEs, including retrotransposons. It is not certain what mechanism has allowed bdelloids to become free of these TEs. Our models provide a plausible scenario for complete elimination of vertically transmitting TEs, assuming that bdelloids have fairly large population sizes and that excision takes place. It is not known what the effective population size, N_e, of bdelloids might be. This can be estimated from measures of sequence diversity, which is a function of $N_e\mu$, where μ is the mutation rate. While the mutation rate of bdelloid rotifers is unknown, there is no evidence that bdelloids have an unusual mutation rate, as relative rate tests between the nucleotide mutation rates of bdelloid and monogonont rotifers show no significant difference at either synonymous or replacement sites (Mark Welch and Meselson 2001). Mitochondrial nucleotide diversity in independently evolving bdelloid clades appears to be similar to that of sexual organisms and averages $\sim$1% (Birky *et al.* 2005; C. W. Birky, personal communication). Mutation rates of mitochondrial DNA are generally higher than those for nuclear DNA, although this varies between taxa. If we suppose a mutation rate of 10^{-9}–10^{-7}, we can estimate the effective population size to be $\sim 10^5$–10^7. This is in the range of being large enough to facilitate complete TE elimination under biologically realistic transposition parameters, even with rates of excision well below that of transposition (see Figures 3 and 4). Thus, unusually high excision rates or attenuated transposition rates are not necessary to explain the existence of TE-free ancient asexuals.

The transposition, excision, and selection parameters that applied to bdelloid rotifers upon the abandonment of sex are also not known. However, it has been suggested that the unusual ecology of bdelloid rotifers, by causing frequent DNA damage and repair, may have facilitated their loss of TEs (M. Meselson, personal communication). Bdelloids are capable of anhydrobiosis, a form of dormancy triggered by desiccation (Ricci 1998). However, unlike most organisms capable of anhydrobiosis, including the monogonont rotifers, bdelloids do not produce the water-replacement substance trehalose (Lapinski and Tunnacliffe 2003). Bdelloids tolerate high levels of ionizing radiation, causing DNA double-strand breaks (DSBs) that are repaired within a few hours (E. Gladyshev and M. Meselson, unpublished results). The bacterium *Deinococcus radiodurans* is well known for its ability to repair DNA damage following radiation or dehydration by efficient repair mechanisms (Battista 1997), suggesting that the response to irradiation in bdelloids may be an outcome of selection for desiccation resistance.

DSB repair processes in bdelloids might either lead to direct elimination of TEs or cause synergistic selection between deleterious TEs, as a result of ectopic exchange between them (M. Meselson, personal communication). Although bdelloids do not undergo meiosis, ectopic pairing of TEs during DSB repair could lead to

deleterious rearrangements or unrepaired breaks, with a frequency proportional to the square of the copy number (Montgomery *et al.* 1987; Charlesworth *et al.* 1994). Such ectopic pairing may lead to the excision of heterozygous TEs, which is thought to be extremely rare for retrotransposons, and would prevent any TE insertion from being selectively neutral, just as in sexual species, enhancing the possibility of complete TE elimination (see Figure 5).

We have shown that newly arisen asexual lineages may be able to avoid early extinction due to TEs. The ability to do so depends critically on large population sizes, with some level of excision. The unique ecology and stress-resistance mechanism of the bdelloid rotifers, which probably evolved as a result of adaptation to their transient aquatic habitats, may have provided the means under which to eliminate deleterious TEs upon the abandonment of sex. Perhaps other asexual lineages were not as fortunate (Arkhipova and Meselson 2005a).

We thank M. Meselson for helpful discussions and L. Loewe for assistance with the simulation modeling. This research was supported by postgraduate scholarships to E.S.D. from the Natural Sciences and Engineering Research Council (Canada) and the University of Edinburgh School of Biological Sciences. B.C. is supported by the Royal Society (United Kingdom).

LITERATURE CITED

Arkhipova, I., and M. Meselson, 2000 Transposable elements in sexual and ancient asexual taxa. Proc. Natl. Acad. Sci. USA **97:** 14473–14477.

Arkhipova, I., and M. Meselson, 2005a Deleterious transposable elements and the extinction of asexuals. BioEssays **27:** 76–85.

Arkhipova, I., and M. Meselson, 2005b Diverse DNA transposons in rotifers of the class Bdelloidea. Proc. Natl. Acad. Sci. USA **102:** 11781–11786.

Arkhipova, I. R., K. I. Pyatkov, M. Meselson and M. B. Evgen'ev, 2003 Retroelements containing introns in diverse invertebrate taxa. Nat. Genet. **33:** 123–124.

Bachtrog, D., and I. Gordo, 2004 Adaptive evolution of asexual populations under Muller's ratchet. Evolution **58:** 1403–1413.

Basten, C. J., and M. E. Moody, 1991 A branching-process model for the evolution of transposable elements incorporating selection. J. Math. Biol. **29:** 743–761.

Battista, J. R., 1997 Against all odds: the survival strategies of *Deinococcus radiodurans.* Annu. Rev. Microbiol. **51:** 203–224.

Birky, Jr., C. W., C. Wolf, H. Maughan, L. Herbertson and E. Henry, 2005 Speciation and selection without sex. Hydrobiologia **546:** 29–45.

Brookfield, J. F. Y., 1991 Models of repression of transposition in *P-M* hybrid dysgenesis by *P* cytotype and by zygotically encoded repressor proteins. Genetics **128:** 471–486.

Brookfield, J. F. Y., and R. M. Badge, 1997 Population genetic models of transposable elements. Genetica **100:** 281–294.

Charlesworth, B., 1985 The population genetics of transposable elements, pp. 213–232 in *Population Genetics and Molecular Evolution*, edited by T. Ohta and K. Aoki. Springer-Verlag, Berlin.

Charlesworth, B., 1990 Mutation-selection balance and the evolutionary advantage of sex and recombination. Genet. Res. **55:** 199–221.

Charlesworth, B., 1991 Transposable elements in natural populations with a mixture of selected and neutral insertion sites. Genet. Res. **57:** 127–135.

Charlesworth, B., and N. H. Barton, 1996 Recombination load associated with selection for increased recombination. Genet. Res. **67:** 27–41.

Charlesworth, B., and D. Charlesworth, 1983 The population dynamics of transposable elements. Genet. Res. **42:** 1–27.

Charlesworth, B., and C. H. Langley, 1986 The evolution of self-regulated transposition of transposable elements. Genetics **112:** 359–383.

Charlesworth, B., and C. H. Langley, 1989 The population genetics of *Drosophila* transposable elements. Annu. Rev. Genet. **23:** 251–287.

Charlesworth, B., A. Lapid and D. Canada, 1992 The distribution of transposable elements within and between chromosomes in a population of *Drosophila melanogaster.* II. Inferences on the nature of selection against element. Genet. Res. **60:** 115–130.

Charlesworth, B., P. Sniegowski and W. Stephan, 1994 The evolutionary dynamics of repetitive DNA in eukaryotes. Nature **371:** 215–220.

Craig, N. L., R. Craigie, M. Gellert and A. M. Lambowitz (Editors), 2002 *Mobile DNA II.* ASM Press, Washington, DC.

Crow, J. F., 1970 Genetic loads and the cost of natural selection, pp. 128–177 in *Mathematical Topics in Population Genetics*, edited by K. Kojima. Springer-Verlag, Berlin.

Docking, T. R., F. E. Saadé, M. C. Elliott and D. J. Schoen, 2006 Retrotransposon sequence variation in four asexual plant species. J. Mol. Evol. **62:** 375–387.

Doolittle, W. F., and C. Sapienza, 1980 Selfish genes, the phenotype paradigm and genome evolution. Nature **284:** 601–603.

Doolittle, W. F., T. B. L. Kirkwood and M. A. H. Dempster, 1984 Selfish DNA and self-restraint. Nature **307:** 501–502.

Edwards, R. J., and J. F. Y. Brookfield, 2003 Transiently beneficial insertions could maintain mobile DNA sequences in variable environments. Mol. Biol. Evol. **20:** 30–37.

Ewens, W. J., 1979 *Mathematical Population Genetics.* Springer-Verlag, Berlin.

Finnegan, D. J., 1992 Transposable elements, pp. 1096–1107 in *The Genome of Drosophila melanogaster*, edited by D. L. Lindsley and G. Zimm. Academic Press, New York.

Hartl, D. L., and S. A. Sawyer, 1988 Why do unrelated insertion sequences occur together in the genome of *Escherichia coli*? Genetics **118:** 537–541.

Hickey, D. A., 1982 Selfish DNA: a sexually-transmitted nuclear parasite. Genetics **101:** 519–531.

Johnson, T., and N. H. Barton, 2002 The effect of deleterious alleles on adaptation in asexual populations. Genetics **162:** 395–411.

Jordan, I. K., L. V. Matyunina and J. F. McDonald, 1999 Evidence for recent horizontal transfer of long terminal repeat retrotransposons. Proc. Natl. Acad. Sci. USA **96:** 12621–12625.

Judson, O. P., and B. B. Normark, 1996 Ancient asexual scandals. Trends Ecol. Evol. **11:** 41–46.

Kaplan, N. L., and J. F. Y. Brookfield, 1983 Transposable elements in Mendelian populations. III. Statistical results. Genetics **104:** 485–495.

Kidwell, M. G., and D. R. Lisch, 2001 Transposable elements, parasitic DNA, and genome evolution. Evolution **55:** 1–24.

Kimura, M., and T. Maruyama, 1966 The mutational load with epistatic gene interactions in fitness. Genetics **54:** 1303–1312.

Langley, C. H., E. A. Montgomery, R. Hudson, N. Kaplan and B. Charlesworth, 1988 On the role of unequal exchange in the containment of transposable element copy number. Genet. Res. **52:** 223–236.

Lapinski, J., and A. Tunnacliffe, 2003 Anhydrobiosis without trehalose in bdelloid rotifers. FEBS Lett. **553:** 387–390.

Malik, H. S., W. D. Burke and T. H. Eickbush, 1999 The age and evolution of non-LTR retrotransposable elements. Mol. Biol. Evol. **16:** 793–805.

Mark Welch, D. B., and M. Meselson, 2001 Rates of nucleotide substitution in sexual and anciently asexual rotifers. Proc. Natl. Acad. Sci. USA **98:** 6720–6724.

Martiel, J.-L., and M. Blot, 2002 Transposable elements and fitness of bacteria. Theor. Popul. Biol. **61:** 509–518.

Maside, X., S. Assimacopoulos and B. Charlesworth, 2000 Rates of movement of transposable elements on the second chromosome of *Drosophila melanogaster.* Genet. Res. **75:** 275–284.

Matsumoto, M., and T. Nishimura, 1998 Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. ACM Trans. Mod. Comput. Simul. **8:** 3–30.

Maynard Smith, J., 1986 Contemplating life without sex. Nature **324:** 300–301.

Montgomery, E. A., B. Charlesworth and C. H. Langley, 1987 A test for the role of natural selection in the stabilization of transposable element copy number in a population of *Drosophila melanogaster.* Genet. Res. **49:** 31–41.

Muller, H. J., 1964 The relation of recombination to mutational advance. Mutat. Res. **1:** 2–9.

Nuzhdin, S. V., 1999 Sure facts, speculations, and open questions about evolution of transposable elements. Genetica **107:** 129–137.

Nuzhdin, S. V., and T. F. C. Mackay, 1995 The genomic rate of transposable element movement in *Drosophila melanogaster.* Mol. Biol. Evol. **12:** 180–181.

Nuzhdin, S. V., and D. A. Petrov, 2003 Transposable elements in clonal lineages: lethal hangover from sex. Biol. J. Linn. Soc. **79:** 33–41.

Nuzhdin, S. V., E. G. Pasyukova, T. V. Morozova and A. J. Flavell, 1998 Quantitative genetic analysis of *copia* retrotransposon activity in inbred *Drosophila melanogaster* lines. Genetics **150:** 755–766.

Orgel, L. E., and F. H. Crick, 1980 Selfish DNA: the ultimate parasite. Nature **284:** 1517–1523.

Pasyukova, E. G., S. V. Nuhzdin and D. A. Filatov, 1998 The relationship between the rate of transposition and transposable element copy number for copia and Doc retrotransposons of *Drosophila melanogaster.* Genet. Res. **72:** 1–11.

Ricci, C., 1998 Anhydrobiotic capabilities of bdelloid rotifers. Hydrobiologia **388:** 321–326.

Sawyer, S. A., D. E. Dykhuizen, R. F. DuBose, L. Green, T. Mutangadura-Mhlanga *et al.*, 1987 Distribution and abundance of insertions sequences among natural isolates of *Escherichia coli.* Genetics **115:** 51–63.

Silva, J. C., E. L. Loreto and J. B. Clark, 2004 Factors that affect the horizontal transfer of transposable elements. Curr. Issues Mol. Biol. **6:** 57–71.

Stephan, W., and Y. Kim, 2002 Recent applications of diffusion theory to population genetics, pp. 72–93 in *Modern Developments in Theoretical Population Genetics*, edited by M. Slatkin and M. Veuille. Oxford University Press, Oxford.

Vieira, C., and C. Biémont, 1997 Transposition rate of the 412 retrotransposable element is independent of copy number in natural populations of *Drosophila simulans.* Mol. Biol. Evol. **14:** 185–188.

Wright, S. I., and D. J. Finnegan, 2001 Genome evolution: sex and the transposable element. Curr. Biol. **11:** R296–R299.

Communicating editor: D. Houle

ORIGINAL ARTICLE

doi:10.1111/j.1558-5646.2007.00118.x

INBREEDING AND OUTBREEDING DEPRESSION IN *CAENORHABDITIS* NEMATODES

Elie S. Dolgin,[1,2] Brian Charlesworth,[1,3] Scott E. Baird,[4] and Asher D. Cutter[1,5]

[1] *Institute of Evolutionary Biology, University of Edinburgh, King's Buildings, Edinburgh, EH9 3JT, United Kingdom*

[2] *E-mail: elie.dolgin@ed.ac.uk*

[3] *E-mail: brian.charlesworth@ed.ac.uk*

[4] *Department of Biological Sciences, Wright State University, Dayton, Ohio 45435*

E-mail: scott.baird@wright.edu

[5] *Department of Ecology and Evolutionary Biology, University of Toronto, 25 Harbord Street, Toronto, Ontario M5S 3G5, Canada*

E-mail: asher.cutter@utoronto.ca

Received November 16, 2006
Accepted February 9, 2007

The nematode *Caenorhabditis elegans* reproduces primarily by self-fertilization of hermaphrodites, yet males are present at low frequencies in natural populations (androdioecy). The ancestral state of *C. elegans* was probably gonochorism (separate males and females), as in its relative *C. remanei.* Males may be maintained in *C. elegans* because outcrossed individuals escape inbreeding depression. The level of inbreeding depression is, however, expected to be low in such a highly selfing species, compared with an outcrosser like *C. remanei.* To investigate these issues, we measured life-history traits in the progeny of inbred versus outcrossed *C. elegans* and *C. remanei* individuals derived from recently isolated natural populations. In addition, we maintained inbred lines of *C. remanei* through 13 generations of full-sibling mating. Highly inbred *C. remanei* showed dramatic reductions in brood size and relative fitness compared to outcrossed individuals, with evidence of both direct genetic and maternal-effect inbreeding depression. This decline in fitness accumulated over time, causing extinction of nearly 90% of inbred lines, with no evidence of purging of deleterious mutations from the remaining lines. In contrast, pure strains of *C. elegans* performed better than crosses between strains, indicating outbreeding depression. The results are discussed in relation to the evolution of androdioecy and the effect of mating system on the level of inbreeding depression.

KEY WORDS: Androdioecy, *Caenorhabditis elegans*, *Caenorhabditis remanei*, hermaphrodite, inbreeding depression, mating system, outcrossing, selfing.

Mating between close relatives causes an increase in homozygosity, which usually results in a decline in fitness known as inbreeding depression. Inbreeding depression is a central factor in the evolution of mating systems, particularly in relation to the evolution of the rate of self-fertilization in hermaphroditic organisms (Jarne and Charlesworth 1993; Uyenoyama et al. 1993). Some models that incorporate inbreeding depression predict the evolution of either complete selfing or complete outcrossing (e.g., Lande and Schemske 1985). Nevertheless, mixed mating systems abound, and other theoretical studies have found conditions under which partial selfing is evolutionarily stable (see review by Goodwillie et al. 2005).

© 2007 The Author(s). Journal compilation © 2007 The Society for the Study of Evolution.

The nematode *Caenorhabditis elegans* exhibits a mating system in which hermaphrodites are self-fertile, but can outcross only with males. This androdioecious reproductive mode makes *C. elegans* a useful system for studying the evolution of outcrossing versus selfing. Under both laboratory and natural conditions, however, males are rare (Hodgkin and Doniach 1997; Barrière and Félix 2005, 2007). Studies of molecular variation in wild-caught strains of *C. elegans* indicate that selfing is the primary mode of reproduction, with low but detectable rates of outcrossing (Denver et al. 2003; Barrière and Félix 2005, 2007; Sivasundar and Hey 2005; Cutter 2006).

There is an ongoing debate about the evolutionary significance of males in *C. elegans*. Although males are present only at extremely low population frequencies, a significant portion of the genome is dedicated to male-specific functions (Jiang et al. 2001), which seems to be maintained by selection (Cutter and Ward 2005). One obvious selective advantage to males is their escape from any inbreeding depression caused by the selfing of hermaphrodites, which is thought to be important in other androdioecious species (Lloyd 1975; Charlesworth 1984; Otto et al. 1993; Rieseberg et al. 1993; Weeks et al. 1999; Stewart and Phillips 2002; Cutter et al. 2003). However, inbreeding depression has not been observed in *C. elegans* for life span (Johnson and Wood 1982; Johnson and Hutchinson 1993), brood size (Chasnov and Chow 2002), and various other life-history traits (Johnson and Hutchinson 1993).

This is consistent with theoretical analyses, which show that very high levels of inbreeding are associated with large reductions in the frequencies of deleterious recessive or partially recessive mutations, leading to reduced inbreeding depression (Charlesworth and Charlesworth 1998). The restriction of gene flow by a predominance of selfing may also lead to greater levels of population subdivision, which could further reduce inbreeding depression (Waller 1993; Theodorou and Couvet 2002; Whitlock 2002; Glémin et al. 2003). In addition, restricted recombination, resulting from low rates of outcrossing and migration (Barrière and Félix 2005, 2007; Cutter 2006), could lead to the accumulation of different favorable combinations of alleles in different local populations, resulting in hybrid breakdown (outbreeding depression) following outcrossing (Templeton 1986).

Phylogenetic evidence indicates that the ancestor of *C. elegans* was gonochoristic (separate males and females)—changes from gonochorism to hermaphroditism have independently occurred at least 10 times in rhabditid nematodes (Kiontke et al. 2004; Kiontke and Fitch 2005). To understand the evolution and regulation of mating systems, comparative studies of inbreeding depression in conspecific and congeneric populations with different modes of reproduction have been undertaken in other groups (e.g., Holtsford and Ellstrand 1990; Demeester 1993; Johnston and Schoen 1996). Given the prominence of *Caenorhabditis* species as model organisms in biological research, it is surprising that little is known about inbreeding depression levels in the genus. Even in *C. elegans*, all the studies mentioned above involved the N2 strain, which has been maintained in the laboratory for thousands of generations. Natural isolates that have not adapted to laboratory conditions might exhibit different levels of inbreeding depression (Stewart and Phillips 2002), especially in view of the observed variation among strains in male production and persistence within laboratory populations (Teotónio et al. 2006).

The purpose of this study was to examine the relationship between breeding system and inbreeding depression, using two *Caenorhabditis* species with contrasting mating systems, *C. elegans* and *C. remanei*. To this end, we used worms recently derived from the wild, and assayed fitness-related traits using similar methodologies for the two species, to obtain estimates of the levels of inbreeding depression. We found that the outcrossing species, *C. remanei*, suffered strong multigenerational inbreeding depression, with the majority of inbred lines going extinct. In contrast, the selfing species, *C. elegans*, mostly exhibited outbreeding depression. We discuss the implications of these results for patterns of genetic diversity and the evolution of mating systems.

Materials and Methods

NEMATODE POPULATIONS

Caenorhabditis elegans strains were recently isolated from three localities in France: Franconville (48°98′N, 2°23′E), Hermanville (49°28′N, 0°32′W) and Merlet (44°45′N, 4°42′E) (Barrière and Félix 2005). These strains were selfed for a few generations to reduce any within-strain heterozygosity before freezing, and were only thawed shortly before initiating the present experiment. Five crosses were established, which represent a variety of within and between population comparisons: three crosses between strains from different populations (JU318 × JU370; JU322 × JU399; JU342 × JU466), one cross between strains isolated within 10 cm of each other in Franconville (JU364 × JU368), and one cross between strains isolated 15 m apart in Merlet (JU314 × JU323). All crossed strains were selected on the criterion of being divergent at a number of amplified fragment length polymorphism (AFLP) markers found among a larger dataset, as described in Barrière and Félix (2005), so as to ensure the crosses were between genetically distinct strains (Table 1). We also present two other measures of genetic divergence from two subsequent studies that used these strains, the number of silent-site pairwise single nucleotide polymorphisms (SNPs) (Cutter 2006) and the mean squared difference in microsatellite repeat length (Barrière and Félix 2007), and test whether genetic distance affects the magnitude of the observed effect.

Table 1. Strains used in crosses, with geographic source in parentheses, and molecular divergence data. The number of differing AFLP markers is from a total of 149 measured and 31 found polymorphic. The number of silent site SNPs is from six regions on two chromosomes spanning a total of 3372.4 silent sites across all loci, with the diversity measure π_{si} shown in parentheses. The mean squared difference in microsatellite repeat length ($\delta\mu^2$) is from six microsatellite loci across five chromosomes.

Strain 1	Strain 2	Geographic distance	Number of differing AFLP markers[1]	Number of silent-site pairwise SNPs (π_{si})[2]	Mean squared difference in microsatellite repeat length[3]
JU364 (Franconville)	JU368 (Franconville)	<10 cm	9	0 (0.0)	0
JU314 (Merlet1)	JU323 (Merlet2)	~15 m	14	21 (0.00614)	28.3
JU318 (Merlet1)	JU370 (Franconville)	570 km	15	5 (0.00144)	3.5
JU342 (Merlet2)	JU366 (Franconville)	570 km	8	21 (0.00610)	21.7
JU322 (Merlet2)	JU399 (Hermanville)	655 km	9	17 (0.00494)	17.8

[1]Barrière and Félix 2005, [2]Cutter 2006, [3]Barrière and Félix 2007.

Males of *C. elegans* were generated in each strain by heat-shock at 26–28°C for ~6 h, and the worms were then returned to 20°C to allow selfing. The resulting male offspring were then crossed in abundance to hermaphrodites, in order to maintain populations with a ~50:50 sex ratio. Each line was then maintained at 20°C for three generations as both a mixed-sex and a pure hermaphrodite population.

Crosses were carried out in four separate experimental blocks at different time points, with two to three genetically distinct crosses per block, and each particular cross was performed twice. It is of interest to note that another cross was also attempted, JU393 × JU407, but the male mating performance of JU393 was so poor that mating frequency was extremely low and these data were discarded. The particularly poor male mating ability of the JU393 strain was also independently observed by H. Teotónio (pers. comm. 2006).

Caenorhabditis remanei populations were established from dauer larvae found living in association with terrestrial isopods in the Wright State University (WSU) Biological Preserve in Southwestern Ohio (39°47′N, 84°03′W). Sampling was done following the protocol of Baird (1999). Briefly, *Trachilipus rathkii* isopods were collected during a single day from several locations within a 0.1 km radius in a wooded area of the WSU Biological Preserve, and sacrificed on 60-mm NGM-lite agar plates seeded with *Escherischia coli* OP50. Nematodes obtained in these collections were considered to be from a single genetically diverse, random mating population, consistent with current knowledge of *C. remanei* from this area (Cutter et al. 2006).

Plates were periodically monitored for ~two days for the appearance of nematodes, and any worms obtained were transferred to a different plate. In total, 30 females and 24 males were isolated, allowed to develop to adulthood, and to mate randomly for ~one day, allowing the possibility of multiple matings, to initiate the study population. The males were then killed and the 30 gravid adult females were subdivided onto three fresh plates, with 10 females per plate, to establish three subpopulations. These plates were then maintained as large outbred populations for ~six generations before initiating the experiment, and concurrently frozen for storage at –80°C. Except where specifically noted, worms of both *C. elegans* and *C. remanei* were maintained at 20°C on 35-mm NGM-lite agar Petri plates seeded with *E. coli* OP50, using standard techniques (Sulston and Hodgkin 1988).

EXPERIMENTAL CROSSES AND LIFE-HISTORY ASSAYS

In *C. elegans*, for each cross we measured fecundity and longevity in four classes of hermaphrodite: pure-strain F1s from each strain, and the two reciprocal hybrid F1s. Individuals of the same class were not independent of each other, as they often came from the same family, considering multiple hermaphrodite offspring of a cross involving a single mother and five to six fathers to belong to the same family. Altogether, 767 worms were assayed, composed of an average of 38 worms per class in each cross, with a mean family size of 5.85, yielding a total number of 178,861 offspring. The general features of the protocol are presented in Figure 1.

We first synchronized populations and cleaned plates of any bacterial contamination by using alkaline hypochlorite (Sulston and Hodgkin 1988). To ensure that the hermaphrodites assayed were the result of crossing rather than selfing, we placed five to six larval stage L4 males from the same strain with one L4 hermaphrodite. After two to three days of mating, each hermaphrodite was moved to a fresh plate to lay eggs for 5 h. After two to three days, we verified that the F1 offspring that developed from these eggs had a male to hermaphrodite ratio that did not differ significantly from 1:1, using a χ^2-test with one degree of freedom. These F1 hermaphrodites were then used for fecundity and longevity assays.

In *C. remanei*, multiple plates from each subpopulation were set up with a single female and a single male at the L4 stage

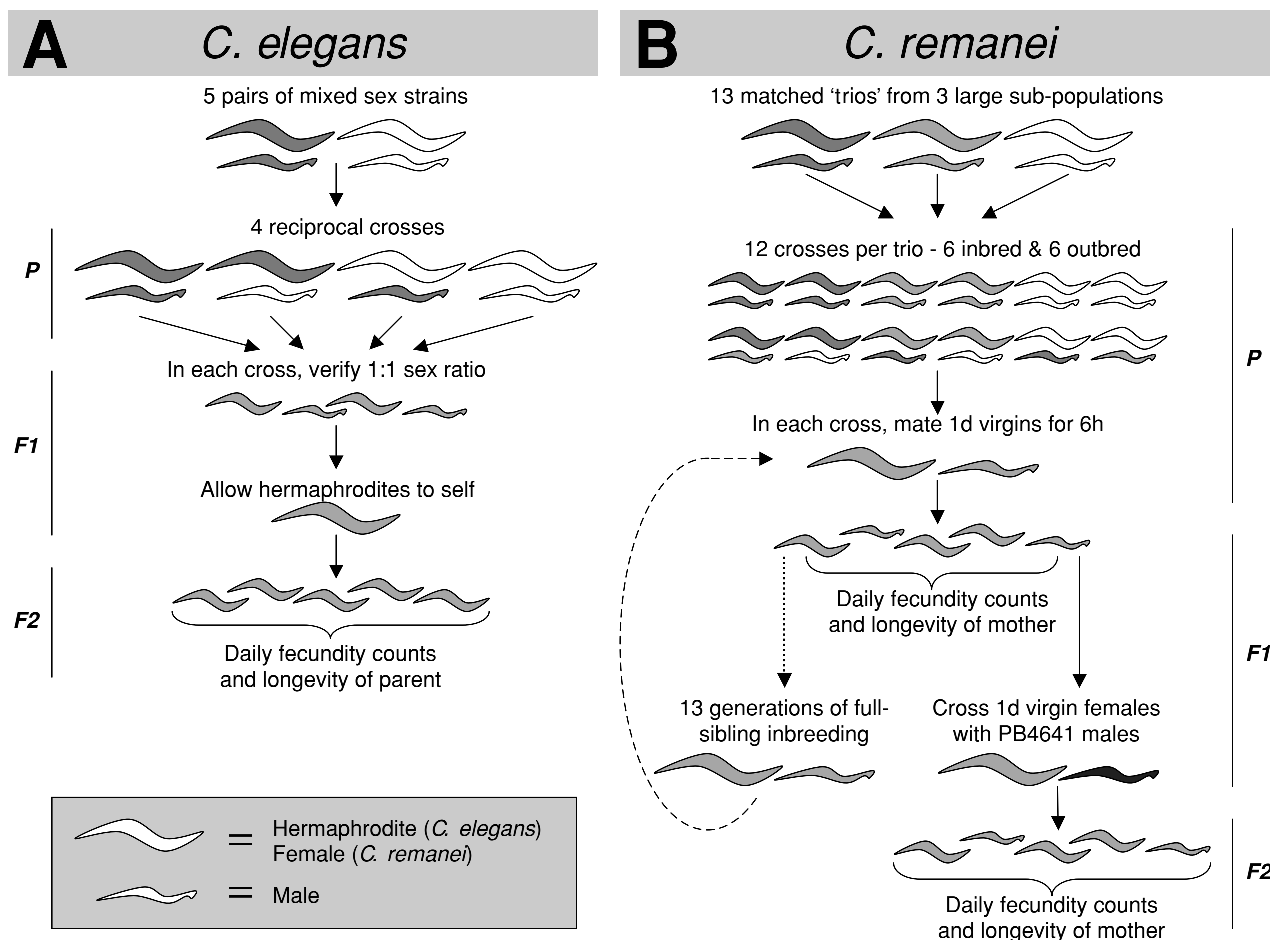


Figure 1. Experimental protocol for *C. elegans* (A) and *C. remanei* (B). The fitness-related traits were daily fecundity, giving a measure of total brood size and relative fitness (w), and longevity. The experimental protocol was designed to measure similar components of fitness for the different breeding systems.

in order to obtain full sibling offspring. Only plates with many offspring (> 300) were subsequently used, to guarantee healthy outbred starting populations. To ensure that the individuals being crossed were as unrelated as possible, we only used crosses between subpopulations for the outbred crosses. This is because individuals within subpopulations may have mated during the ~six generations between the field collection and establishment of the subpopulations, and initiating the experiment. To this end, we randomly chose one plate from each of the three subpopulations to create 13 "trios."

Within each trio of subpopulations, we set up six inbred crosses, two from each subpopulation, and all six possible outbred crosses between each subpopulation, including both the reciprocal crosses (Fig. 1B). We then measured productivity and longevity in the resulting outbred and full-sibling inbred females, and in the F1 progeny of these individuals crossed to males of the standard strain, PB4641 (selected because it is a highly inbred strain, which is being used for genomic sequencing). Using PB4641 males allows us to standardize the male contribution in all treatments and isolate the genetic effects of inbreeding depression on the F1 females. A total of 215 worms were assayed, 71 inbred and 63 outbred individuals, of which 46 inbred F1 and 35 outbred F1 progeny were crossed with PB4641. A total of 60,750 offspring were counted.

In addition, we maintained a number of inbred lines for 13 generations of full-sibling mating. We attempted to propagate 39 randomly selected lines, representing all three subpopulations taken from inbred F1 progeny, by placing full-sibling single males with single females, isolated as L4s. In establishing our inbred lines, we set up two mating plates of each inbred line in each generation, to allow for failure to mate. If no progeny were found on either plate, the line was deemed to have gone extinct. Otherwise, one of the plates with progeny was selected at random to propagate the inbred line. After 13 generations, we measured the fecundities of six classes of individuals from the extant inbred lines: (1) full-sibling inbred females, (2) inbred females crossbred with males from different inbred lines, (3) outbred females thawed from the three subpopulations frozen before initiation of

the experiment, (4) inbred females crossed to PB4641 males, (5) crossbred F1 females whose parents were from different inbred lines crossed to PB4641 males, and (6) outbred females crossed to PB4641 males. At this stage, a further 170 worms were assayed and 38,106 offspring counted. We also froze the extant inbred lines at –80°C and repeated these productivity measures again in a second block.

Fecundity and longevity assay methods were similar between the two species, with a few notable exceptions. In *C. elegans,* L4 hermaphrodites were isolated onto individual plates (there was no discernible difference in the timing of development between pure-strain and hybrid worms). Once they had developed into adults, they were allowed to lay eggs, and were transferred to fresh plates every 24 h for three days, and then to new plates to measure late fecundity and longevity. In *C. remanei*, male and female virgins were isolated at the L4 stage and kept separate for 40–44 h until they had become ∼ one-day-old adults. Although serially inbred lines of *C. remanei* had delayed growth rates, we ensured that they had molted to adults many hours before the time of mating. Female reproduction in *C. remanei* is limited by the number of sperm transferred upon mating. Consequently, we placed a single male and female together on a fresh plate and permitted them to mate for 6 h to mirror the sperm-limitation experienced by selfing protandrous *C. elegans* hermaphrodites, where spermatogenesis precedes oogenesis, and once egg maturation begins, the hermaphrodite possesses its full complement of sperm. The male was then killed and the female moved to a new plate for 18 h, transferred every 24 h for two more days, and then moved to new plates to measure late fecundity and longevity. Plates without any offspring were observed in 13 of the 156 crosses in the parental generation, with no difference between inbred and outbred crosses ($\chi^2_1 = 0.08$, $P = 0.78$). It was assumed that mating had not been successful on these plates, and they were omitted from further analyses.

For both species, after the timed plate transfers for days 1–3, the worms were transferred after two days, and subsequently only if offspring or bacterial contamination was observed, although these were both rare events. Worms were then checked every one to two days, and were deemed to have died if they did not respond to gentle agitation with a platinum pick. Eggs on the laying plates were allowed to hatch and develop for three days, and were then counted to give daily egg-to-late-larval fecundities for days 1–3. For *C. remanei*, day 1 was divided between the first 6 h mating period and the subsequent 18 h, which was denoted as day 1.5 for the calculation of w (see below). Late fecundity beyond day 3 was summed and included in measures of total brood size.

Daily fecundities were used to generate two fitness components: brood size (the unweighted sum of all progeny over the entire life span including late productivity) and relative fitness, w. This second measure is proportional to the expected fitness of an age-structured population, defined as $w = \sum_x e^{-rx} l_x m_x$ (Charlesworth 1994), where $l_x m_x$ is the product of survivorship to age x and fecundity at day x, and r is a constant scaling term, equal to the growth rate of the population as a whole. Due to the short development time of these worms and the sperm-limited fecundity of *C. elegans* hermaphrodites or singly mated *C. remanei* females, the timing of egg laying is the principal factor affecting w. For *C. elegans*, r was calculated by defining the grand mean fitness of the pure strains as $\bar{w}_{\text{pure-strain}} = 1$, obtained using $l_x m_x$ pooled across all pure-strain worms. This r value was used to weight the productivity by a negative exponential function of age. For *C. remanei*, r was calculated by defining the mean fitness of the outbred individuals in the parental cross as $\bar{w}_{P-\text{outbred}} = 1$. This value of r was then used to calculate w for the crosses with PB4641 males and for the crosses after 13 generations of inbreeding.

STATISTICAL ANALYSIS

Longevity was analyzed by Kruskal–Wallis tests, and brood size and w were analyzed by general linear mixed models, using the JMP statistical package, version 5.1 (SAS Institute, Cary, NC). Residuals were normally distributed for brood size and w in *C. elegans*, and for w in *C. remanei*, but brood size data for *C. remanei* had to be square-root-transformed to normalize the residuals, because of the large differences between inbred and outbred worms. For *C. elegans*, fixed factors were breeding class (i.e., hybrid vs. pure strain), cross identity, block, block × breeding class, block × cross identity, and breeding class × cross identity. A random effect in the analysis was family, with Z-statistics used to test significance. Heterosis was also measured within each cross. Heterosis is usually defined as the relative increase in fitness of hybrids between strains due to increased heterozygosity. We measure it as the excess of the mean fitness of the hybrid F1s, relative to the mean of the pure-strain values (Falconer and Mackay 1996, p. 253), defined here as

$$H_j = \left[\frac{\bar{X}_j(\text{hybrid}) - \bar{X}_j(\text{purestrain})}{\bar{X}_j(\text{purestrain})}\right],$$

where $\bar{X}_j$(hybrid) and $\bar{X}_j$(purestrain) represent the means for the particular trait under investigation, X, of hybrid and pure-strain F1s within the jth cross. Thus, a positive value of H denotes inbreeding depression, and a negative value of H indicates outbreeding depression. The means within the same cross were compared using t-tests.

For *C. remanei*, inbreeding level and trios of matched subpopulations were included as fixed factors in the analysis of brood size and w for the assays of the first parental (P) generation of inbreeding and the F1 cross with PB4641 males. For the P generation cross, the inbreeding levels were sibling-mated and randomly mated, and for the F1 × PB4641 cross the inbreeding levels were inbred and outbred. For the assays after

13 generations of inbreeding, brood size and w were measured, but not longevity. Two general linear models were analyzed, with block and inbreeding class included as fixed factors. In the first analysis, the inbreeding classes were inbred (within lines), crossbred (the crosses between different inbred lines) and outbred, while in the second, the F1s of these classes crossed to PB4641 males were considered. The means were compared between the fully inbred, crossbred, and outbred females and the F1s crossed with PB4641 males, using Tukey tests. To analyze the rate of extinction of inbred lines over the course of 13 generations of serial inbreeding, we performed random permutation tests for a least-squares linear regression model, in order to determine appropriate significance thresholds that take into account the correlations among extinction rates between successive generations. The permutation tests were performed by holding the generation numbers constant and permuting the per generation rates of extinction 10^7 times, and then comparing the resulting r^2 values.

Results

OUTBREEDING DEPRESSION IN *C. ELEGANS*

For both total brood size and relative fitness (w), pure-strain F1s performed significantly better than hybrids ($F_{1,631} = 16.4$, $P < 0.0001$ for brood size; $F_{1,631} = 22.7$, $P < 0.0001$ for w). Both of these traits were affected by the cross identity ($F_{3,631} = 10.9$, $P < 0.0001$ for brood size; $F_{3,631} = 8.3$, $P < 0.0001$ for w), and the magnitude of outbreeding depression also varied significantly between different crosses (as indicated by breeding class × cross identity: $F_{4,631} = 8.5$, $P < 0.0001$ for brood size; $F_{4,631} = 8.3$, $P < 0.0001$ for w). The overall trend was significantly towards outbreeding depression with mean heterosis values of –0.064 and –0.093 for brood size and w, respectively (Fig. 2). When examining each cross independently, however, outbreeding depression was only significant in some crosses (three of five for brood size; two of five for w), and one cross showed a significant positive heterosis value for brood size (Fig. 2).

The block in which assays were carried out was also a significant factor for brood size ($F_{2,631} = 11.1$, $P < 0.0001$), but not for w ($F_{4,631} = 2.81$, $P = 0.06$). Interactions between block and genetic factors for brood size were not significant ($F_{3,626} = 0.64$, $P = 0.59$ for block × cross identity; $F_{3,626} = 2.1$, $P = 0.13$ for block × breeding class). This indicates that experimental noise was introduced into the data by undertaking the experiment over four blocks. However, the lack of significant interactions or strong block effects on w suggests that these effects were, for the most part, in the same direction for all genotypes.

Relative fitness, *w,* is measured from daily progeny counts and so is not independent of total brood size. These two measures were significantly positively correlated ($r = 0.475$; $t_{764} = 13.1$, $P < 0.0001$), with the magnitude and direction of outbreeding depression for both brood size and w in the same direction ($r = 0.88$), although this is limited by sample size ($t_3 = 1.31$, $P = 0.14$). Neither geographic proximity nor the number of polymorphic AFLP markers shared between the crossed strains affected the direction of outbreeding depression ($P > 0.59$ for all regression analyses). There was, however, a nonsignificant trend towards a positive relationship between the magnitude of outbreeding depression and both the number of pairwise single-nucleotide polymorphism differences ($F_{1,3} = 4.49$, $P = 0.12$ for brood size; $F_{1,3} = 2.06$, $P = 0.25$ for w) and the mean squared difference in microsatellite repeat length ($F_{1,3} = 12.46$, $P = 0.03$ for brood size; $F_{1,3} = 6.98$, $P = 0.07$ for w) (see Table 1). This suggests that more divergent strains may exhibit stronger outbreeding depression. More crosses would be needed to test this possibility.

Unlike for total brood size and w, hybrids and pure-strain worms showed no difference in longevity ($\chi^2_1 = 0.12$, $P = 0.73$). Although the five crosses exhibited different life spans($\chi^2_4 = 20.7$, $P = 0.0004$), there was no difference in the longevity of pure-strain versus hybrid F1s in any of the individual crosses (all $P > 0.05$; see Fig. 2C). Blocks also had significant effects on longevity ($\chi^2_3 = 35.1$, $P < 0.0001$), although breeding class remained nonsignificant when considering any particular block independently (all $P > 0.05$). Family was a significant effect for brood size ($Z = 4.8$, $P < 0.0001$), w ($Z = 5.2$, $P < 0.0001$), and longevity ($Z = 1.9$, $P = 0.03$), suggesting a considerable amount of variation among parents of the same genotype. This is probably not due to genetic diversity in the parental strains, since the strains used here were selfed for a few generations before freezing, and heterozygosity was found to be quite low in a subsample of fresh isolates collected from the same locations as the ones being used here (Barrière and Félix 2005). It is more likely that this is caused by variation in the ages of the mothers. Eggs laid by older worms tend to be more developed than those laid by younger worms. Therefore, even though we used a timed egg lay, the ages may not have been developmentally synchronized if the mothers were of different ages (Peters et al. 2003).

INBREEDING DEPRESSION IN *C. REMANEI*

Full-sibling mated females had a lower total brood size and w compared with randomly mated females, although this effect was not significant ($F_{1,120} = 1.43$, $P = 0.23$ for brood size; $F_{1,120} = 1.35$, $P = 0.25$ for w). However, when these females' F1 progeny were crossed with PB4641 males, inbreeding level was a significant factor for both brood size ($F_{1,67} = 4.02$, $P = 0.049$) and w ($F_{1,67} = 7.77$, $P = 0.007$) (Fig. 3). Our experimental design allows us to try to disentangle the sources of inbreeding depression. The poorer performance of sibling-mated females in the parental generation, although nonsignificant (probably because of high variance in the productivity-related traits), can possibly be attributed to embryonic or larval inviability. The difference in

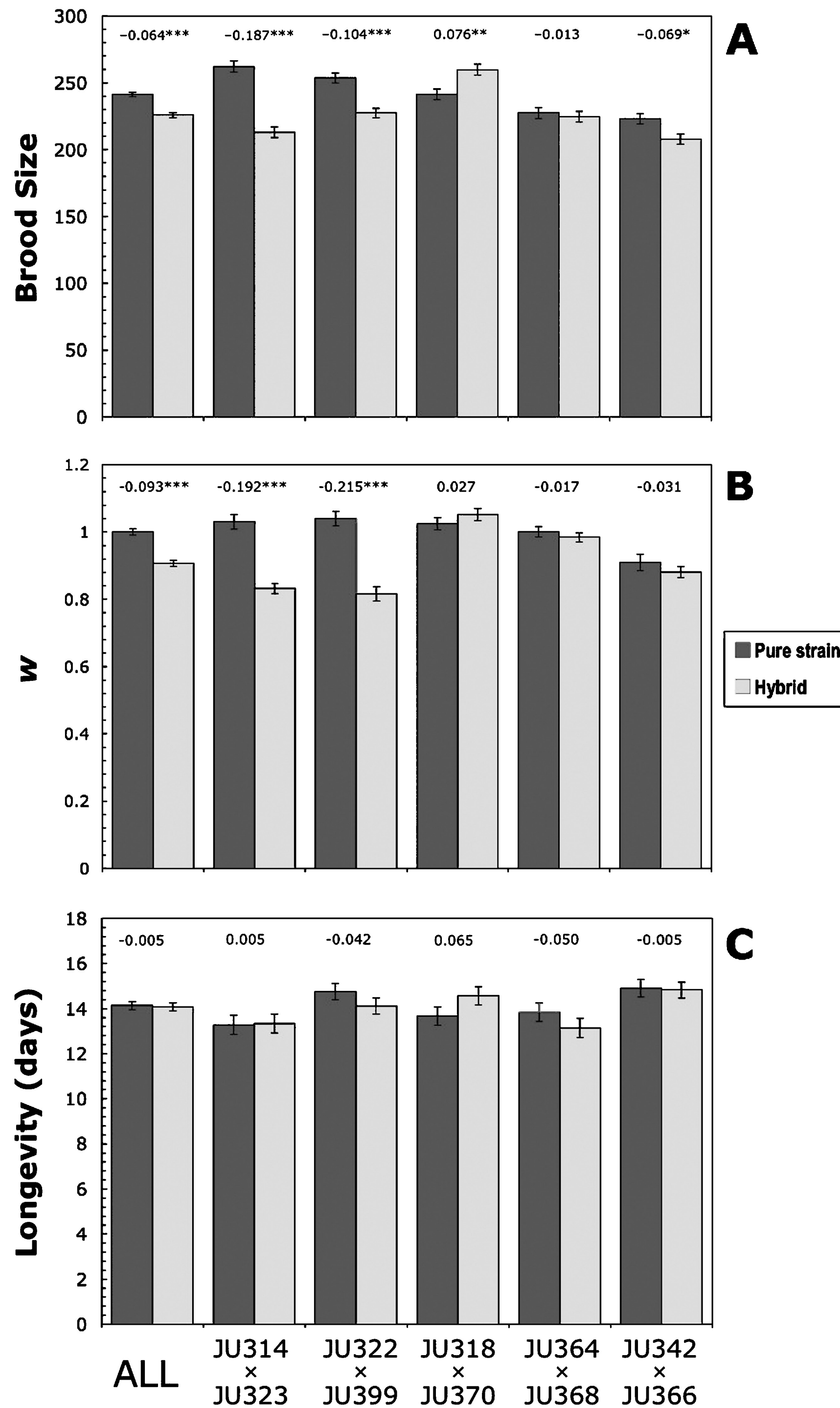


Figure 2. **Measurements of mean pure-strain (darkly shaded bars) and hybrid (lightly shaded bars) brood sizes (A), relative fitnesses (B), and longevities (C) by cross identity and overall in *C. elegans*. Numbers above the bars show heterosis values, calculated as (hybrid/pure strain)—1. Asterisks indicate significance of difference between pure-strains and hybrids in a Student's *t*-test. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. Error bars indicate ± 1 SE.**

the F1s crossed to PB4641 males indicates reduced egg production and/or a reduction in the timing of egg laying in the inbred females, since the number of sperm and the mating performance of the PB4641 fathers are invariant. In both generations among crosses, inbreeding level had no effect on longevity ($\chi^2_1 = 0.01$, $P = 0.91$ for parentals; $\chi^2_1 = 0.09$, $P = 0.77$ for F1 × PB4641). There were no significant effects of the trios of randomly matched subpopulations on any of the measures in either of the crosses.

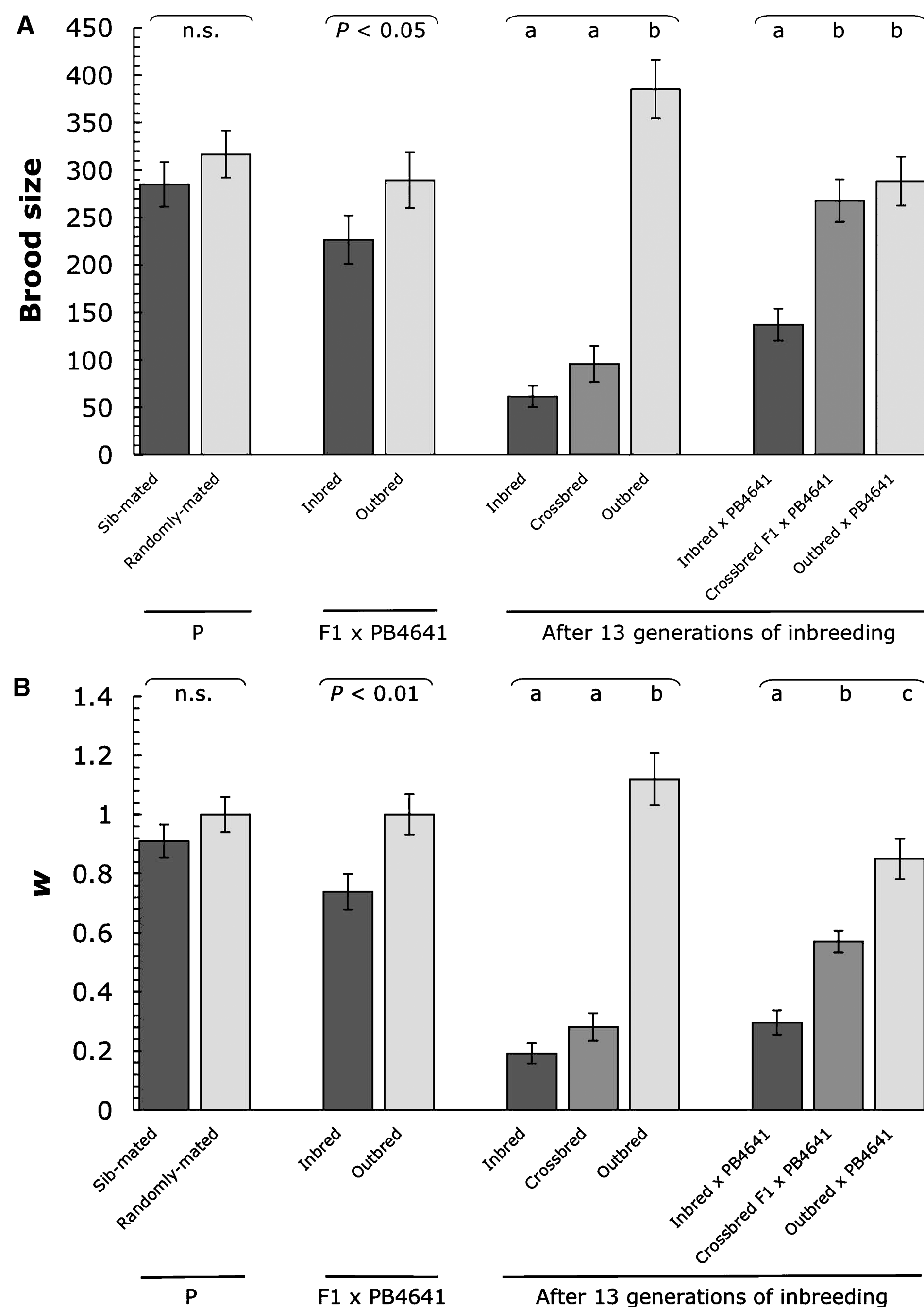


Figure 3. **Measurements of brood size (A) and relative fitness (B) in *C. remanei* for the parental generation (P), the F1 females mated to PB4641 males (F1 × PB4641), and crosses after 13 generations of full-sibling mating. Darkly shaded bars denote sibling-mated females and their inbred progeny, lightly shaded bars denote randomly mated females and their outbred progeny, and intermediate shaded bars denote inbred females crossbred with males from different inbred lines and their progeny. Post hoc significance tests were done separately for each of the four sets of crosses—Student's *t*-tests for *P* and F1 × PB4641, and two separate Tukey tests for crosses after 13 generations of inbreeding. Each comparison is represented by a curved line above the bars being evaluated with the significance denoted underneath. Identical letters within a comparison indicate nonsignificant differences, but letters across comparisons or across panels are unrelated. Error bars indicate ± 1 SE.**

For the assay of total brood size and w on the five extant lines after 13 generations of full-sibling inbreeding, the effects of inbreeding level were highly significant ($F_{2,72} = 51.3$, $P < 0.0001$ for brood size; $F_{2,72} = 56.4$, $P < 0.0001$ for w). Outbred females had greater brood size and relative fitness than inbred females mated to males from the same ("inbred") or different ("crossbred") inbred lines (Fig. 3). Crossbred females had ~ 50% greater mean performance than inbred females, although this effect was not

statistically significant due to large variances ($q_{72,3} = 1.67$, $P > 0.20$ for brood size; $q_{72,3} = 1.16$, $P > 0.50$ for w). The lower values for crossbred versus outbred females suggests that these inbred parents have reduced gamete production and/or mating ability, although the increased performance of crossbred versus inbred females suggests that larval inviability may play a role in the source of inbreeding depression as well.

Taking the F1 progeny females from the crosses between the strains surviving after 13 generations of full-sibling inbreeding and mating them to PB4641 males also yielded a highly significant effect of inbreeding level on fitness ($F_{2,92} = 12.3$, $P < 0.0001$ for brood size; $F_{2,92} = 29.5$, $P < 0.0001$ for w). Outbred females crossed to PB4641 males had greater trait values than inbred females crossed to PB4641 males (Fig. 3), with the reduced performance of the inbred females being significantly greater after 13 generations of inbreeding as compared to after 1 generation ($t_{70} = 2.91$, $P = 0.005$ for brood size; $t_{70} = 5.95$, $P < 0.0001$ for w).

Crossbred F1 females mated with PB4641 males performed significantly better than inbred females ($q_{92,3} = 6.00$, $P < 0.001$ for brood size; $q_{92,3} = 6.21$, $P < 0.001$ for w); however, despite no difference in the mean total brood sizes for crossbred F1 females and outbred females mated to PB4641 males ($q_{92,3} = 1.10$, $P > 0.50$), w was significantly lower for crossbred F1 females ($q_{92,3} = 5.92$, $P < 0.001$). Because crossbred F1 individuals are the progeny of unrelated inbred lines, they have an inbreeding coefficient of 0. The near equivalence of brood sizes indicates that crossbred F1 females had fecundity levels restored to outbred levels, but the reduction in w suggests a maternal effect on the timing of egg production, because the parental lines of the crossbred F1 female had been inbred for 13 generations. There was little evidence of purging of deleterious mutations during the process of inbreeding, because the crossbred F1 females did not perform significantly better than outbred females for either total brood size or w (Crnokrak and Barrett 2002). Block effects were not significant for any of the traits in *C. remanei*.

Over the course of serially inbreeding for 13 generations, 34 of the 39 lines went extinct. To avoid the possibility of lines failing to mate successfully, we set up two plates of each inbred line in each generation. From the first generation of inbreeding, we found that 8.3% of single male–female crosses failed to mate in a 6 h period. Because we observed no difference between inbred and outbred crosses in the frequency of failure to mate in both the parental cross ($\chi^2_1 = 0.08$, $P = 0.78$) and the F1 × PB4641 cross ($\chi^2_1 = 0.05$, $P = 0.83$), we assume that failure is a random event; therefore, < 1% of lines should fail to mate on both plates, although this is conservative because we allowed the inbred lines to mate indefinitely. Figure 4 shows the per generation and cumulative rates of extinction of inbred lines. The per generation extinction rate was consistently much greater than 1% ($t_{12} = 4.00$, $P = 0.0009$), with a mean of 13.9% (SE = 3.2%), and there was a positive association between the per generation rate of extinction and time, as indicated by a permutation test on a linear regression ($r^2 = 0.44$, $P = 0.015$). Because this significance value is nearly equivalent to that found for a parametric regression analysis ($t_{11} = 2.97$, slope = 0.020, $P = 0.013$), we used a parametric quadratic regression to test for purging. Purging of deleterious mutations over time would be expected to lead to a decreasing rate of extinction (Crnokrak and Barrett 2002); however, a quadratic regression did not provide a better fit to the data ($t_{10} = 0.51$, $P = 0.62$), indicating that fitness decline does not taper off. It is worth noting that the extant lines after 13 generations were derived from all three initial subpopulations. Preliminary data from sequencing of exonic nuclear DNA indicates a large number of single nucleotide

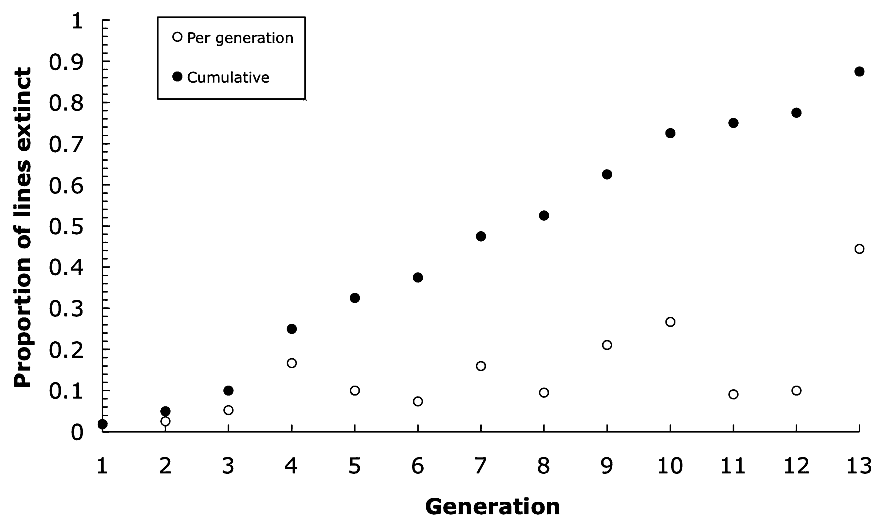


Figure 4. The rate of extinction of inbred lines in 13 generations of serial inbreeding. The open circles show the per generation proportion of lines going extinct; the closed circles show the cumulative proportion.

polymorphisms between the inbred lines, with some residual heterozygosity in at least one of the lines (results not shown). Thus, there appears to be no tendency to fix particular genotypes from a single subpopulations.

Discussion

OUTBREEDING DEPRESSION IN *C. ELEGANS*

Outbreeding depression is defined as the reduction in fitness following hybridization between divergent populations (Templeton 1986). When this occurs in first generation (F1) hybrids, as in our study of *C. elegans*, it could be caused by a disruption of beneficial interactions at three possible levels: between genes and the environment (local adaptation), within loci (underdominance), and between loci (coadapted gene complexes). The first two possibilities can probably be ruled out in our case. First, any effects of adaptation to local environmental conditions should be absent in the constant artificial laboratory environment, and there is no reason to suppose that outbred individuals should be any less adapted to this novel environment. Second, it is well known that underdominance does not lead to the maintenance of variability, even with high selfing rates (Kimura and Ohta 1971; Rocheleau and Lessard 2000), unless there is frequency-dependent selection under conditions similar to those involving local adaptation (Wilson and Turelli 1986). Therefore, the observed pattern of outbreeding depression on fecundity can probably be attributed to a breakdown of coadapted gene complexes, revealing deleterious allele combinations affecting the fitness of the F1 hybrids or the viability of the F2 recombinant progeny (Phillips and Johnson 1998). On the other hand, we did not find any inbreeding or outbreeding depression for longevity. This is consistent with earlier studies that found no effect of breeding class on life span (Johnson and Wood 1982; Johnson and Hutchinson 1993), and little evidence of directional effects of new mutations on longevity (Keightley and Caballero 1997; Vassilieva and Lynch 1999; Keightley et al. 2000; Halligan et al. 2003).

We studied a limited number of crosses between strains from France; however, because *C. elegans* shows little or no geographic structuring of molecular diversity on a large scale (Barrière and Félix 2005, 2007; Haber et al. 2005; Cutter 2006), outbreeding depression is probably not limited to strains from France. Indeed, outbreeding depression appears to be a global phenomenon in *C. elegans*. In crosses between the canonical N2 strain and a Hawaiian strain, CB4856, the progeny of F1 hybrids exhibit a significant increase in embryonic and early larval lethality, with similar hybrid incompatibilities observed between other wild strains (M. Rockman, H. Seidel, and L. Kruglyak, pers. comm. 2006; M. Ailion, pers. comm. 2006). In the related selfing species, *C. briggsae*, reproductive isolation was also found in crosses between strains AF16 and HK104, with approximately one third of F2 progeny exhibiting a delay in development and reduced intrinsic growth rates (S. Baird, unpubl. results).

Previous studies failed to show any effect of inbreeding level on fitness-related traits in *C. elegans* (Johnson and Wood 1982; Johnson and Hutchinson 1993; Chasnov and Chow 2002). However, as pointed out in the introduction to this paper, these all involved crosses where one of the strains was the long-maintained N2 strain, and the partner strains in the crosses had also been in culture for extended periods of time (Hodgkin and Doniach 1997). Here we attempted to overcome this problem by using strains recently caught in the wild. Using similar laboratory techniques to previous studies, we found a significant reduction in the performance of outcrossed individuals. Obtaining estimates under standardized optimal conditions is an important first step, but it is uncertain how this effect translates to performance under natural conditions. It is generally assumed that inbreeding depression is enhanced in harsher natural environments (see reviews by Crnokrak and Roff 1999; Keller and Waller 2002; Armbruster and Reed 2005). Whether this is also widely true of outbreeding depression remains to be seen, although in the selfing hermaphroditic snail, *Physa acuta*, outbreeding depression was observed only in the field but not in the lab (Henry et al. 2003). Potential evidence of outbreeding depression and selection against hybrids in wild populations of *C. elegans* comes from recent regular monitoring of haplotype frequencies in a population in France, where the frequency of a recombinant haplotype decreased significantly over the span of a year (Barrière and Félix 2007). Because the ecology of *C. elegans* is poorly understood, data from more natural conditions are needed.

IMPLICATIONS FOR *C. ELEGANS* DIVERSITY

It has long been recognized that selfing can promote the maintenance of linkage disequilibrium among loci whose alleles have epistatic effects on fitness ("coadaptation") (e.g., Stebbins 1957; Allard 1975), although selfing can also enhance randomly generated linkage disequilibrium between neutral variants, because it reduces the effective rate of recombination (Charlesworth 2003). In *C. elegans*, extremely low levels of diversity are found, with extensive linkage disequilibrium across the genome, both within and between chromosomes (Koch et al. 2000; Barrière and Félix 2005, 2007; Haber et al. 2005; Cutter 2006). This pattern is thought to be determined largely by self-fertilization and population subdivision, combined with rare but regular migration and outcrossing.

Our results suggest that outbreeding depression could be an additional important factor in shaping the genetic structure of *C. elegans*. Selection against hybrids would result in a reduction in the effective migration and outcrossing rates (Barton and Bengtsson 1986), thereby maintaining higher levels of linkage disequilibrium than in its absence. Outbreeding depression may also help to explain the large discrepancies between estimates of

outcrossing rates based on heterozygote frequencies at microsatellite loci (Barrière and Félix 2005, 2007; Sivasundar and Hey 2005), and measures from linkage disequilibrium, which are typically one to three orders of magnitude lower (Barrière and Félix 2005, 2007; Cutter 2006). If selection maintaining coadapted gene complexes caused hybrids between strains to have reduced contribution to the next generation, then the effective outcrossing rate will be greatly reduced.

FREQUENCY OF MALES

Because males are nonessential for reproduction in *C. elegans*, their maintenance requires explanation. It is possible that androdioecy could simply be a byproduct of the sex-determination mechanism, which allows male production via nondisjunction of X-chromosomes in hermaphrodite meiosis (hermaphrodites and females are XX, males are XO). Chasnov and Chow (2002) argued that males are nonadaptive but persist because mating is sufficiently frequent to preclude degeneration of male-specific genes by deleterious mutation. Other arguments favor a selective advantage of outcrossing to explain the persistence of males, and it is conceivable that males and outcrossing are advantageous under different ecological settings (Schulenburg and Müller 2004), or particular demographic scenarios involving metapopulation dynamics (Pannell 2002; Weeks et al. 2006a). Our results, however, suggest that males and outcrossing will be selected against in natural populations, because the existence of outbreeding depression must reduce the frequency of males below that expected in its absence. This was verified by using equation (3a) of Cutter et al. (2003). For example, with the male reproductive efficiency calculated for the N2 strain (Cutter et al. 2003), which has a fairly typical male mating ability for the species (Teotónio et al. 2006), the observed level of outbreeding depression of 6–9% would reduce the expected equilibrium frequency of males by about one-third for any rate of X chromosome nondisjunction.

C. *REMANEI* INBREEDING DEPRESSION

Consistent with expectations (e.g., Lande and Schemske 1985; Charlesworth and Charlesworth 1998), the outcrossing species *C. remanei* exhibited much greater levels of inbreeding depression than the selfing *C. elegans*. This appears to reflect the effects of inbreeding on various components of fitness. *Caenorhabditis remanei* female reproduction is limited by the number of sperm transferred by their male partners, which is a product of sperm production and mating efficiency. We attempted to standardize the male contribution by allowing mating for a controlled period of time. We also ensured that mating had indeed occurred by discarding any crosses for which no progeny were observed, although this might underestimate inbreeding depression by ignoring crosses that were unsuccessful due to genetic incompatibilities rather than a failure to copulate. Also, by mating the F1 female progeny to PB4641 males, sperm limitation and male mating efficiency should be constant and independent of inbreeding level. This allowed us to compare directly the performance of inbred and outbred females. In these crosses, inbreeding depression will be solely due to females, whereas without the control PB4641 males, inbreeding depression can also be affected by any male component and by larval inviability of offspring. Comparisons of parental crosses with F1 × PB4641 crosses indicate that inbreeding depression in *C. remanei* is affected by both male and female performance, in addition to larval inviability (see Fig. 3).

EVIDENCE OF MATERNAL-EFFECT INBREEDING DEPRESSION

The result that crossbred F1 females mated to PB4641 males had lower w than outbred females, despite having nearly the same brood sizes, suggests that outcrossing does not completely eliminate the effect of prior inbreeding on fitness in *C. remanei*. This indicates a maternal effect on inbreeding depression. A few other studies have found that maternal inbreeding can influence outcrossed progeny fitness, although the magnitude of these effects is generally less than the direct genetic effects of inbreeding depression (e.g., Hauser and Loeschcke 1995; Lyons 1996; Vogler et al. 1999; Hayes et al. 2005). This highlights the need to separate the sources of inbreeding depression, in order to accurately estimate the magnitude of the direct genetic effects. Our initial F1 × PB4641 crosses should not have been influenced by maternal inbreeding depression, since the parents were outbred, so these should provide a good estimate of zygotic inbreeding depression.

The significantly lower w suggests that it is the timing of egg production and fertilization rather than the absolute brood size that is altered by maternal-effect inbreeding depression. In *C. elegans*, a number of maternal-effect genes have been identified that influence developmental and behavioral timing, known as Clk genes (Lakowski and Hekimi 1996). For example, *clk-1* mutants have a two- to six-fold reduction in their egg production rate, but display maternal rescue, indicative of maternal effects (Wong et al. 1995). It is possible that some of the inbred *C. remanei* lines in our experiment became homozygous for genes with similar effects, thereby causing maternal-effect inbreeding depression on physiological traits related to the timing of the reproductive schedule in the crossbred F1 progeny.

EVOLUTION OF ANDRODIOECY

Recent evidence suggests that the genetics of switching between different reproductive modes is relatively simple in *Caenorhabditis* species, requiring a very limited number of mutations (Nayak et al. 2005; Braendle and Félix 2006; Hill et al. 2006). This demonstrates the potential for frequent evolutionary transitions in mating systems in *Caenorhabditis*. Selfing could have many advantages in *Caenorhabditis*, most notably reproductive assurance in species

colonizing ephemeral habitats with fluctuating population dynamics and episodic low densities, although the natural history of the genus is poorly understood and it is unclear exactly how the habitats differ between the gonochoristic and androdioecious species (Kiontke and Sudhaus 2006; Weeks et al. 2006a). The potential benefits of selfing, however, must be weighed against the cost of overcoming the strong inbreeding depression found in *C. remanei*, which presumably reflects conditions experienced by a lineage upon first shifting to hermaphroditism.

Over the course of 13 generations of full-sibling mating in *C. remanei*, 87% of our inbred lines went extinct and the surviving lines were extremely sick, with very low fecundity, and exhibited both direct and maternal-effect inbreeding depression (see Figs. 3 and 4). Among the extant lines, there was little evidence of any rebound in fitness over time that would have indicated purging of deleterious alleles (Crnokrak and Barrett 2002), which suggests that the inbreeding depression is largely due to the cumulative effects of many deleterious alleles of small effect rather than a few segregating deleterious alleles of large effect (Willis 1999). Indeed, most inbred strains of *C. remanei* die out in the laboratory, and those that survive have much lower fecundity and longer generation times than noninbred strains, and appear to be less fit than recently established strains (S. Baird and E. Dolgin, unpubl. observations). Furthermore, newly arising hermaphroditic lineages would probably have suboptimal sperm numbers and brood sizes, thereby decreasing the rate of population expansion (Cutter 2004), lowering the effective population size, and increasing the likelihood of extinction.

In the long run, selfing lineages that survive are likely to be those that have purged much of their genetic loads; a higher degree of population subdivision in selfing lineages may facilitate such purging (Waller 1993; Theodorou and Couvet 2002; Whitlock 2002; Glémin et al. 2003). This makes it hard for transitions back to outcrossing to occur (Charlesworth and Charlesworth 1998). But are transitions to hermaphroditism in rhabditid nematodes evolutionary dead ends? Phylogenetic evidence suggests that hermaphroditism has evolved frequently in rhabditids, but these lineages rarely survive in the long run (Kiontke and Fitch 2005). This is consistent with the pattern found in several plant genera (Schoen et al. 1997; Goodwillie 1999; Truyens et al. 2005; but see Takebayashi and Morrell 2001). On the other hand, in the clam shrimp *Eulimnadia*, androdioecy appears to be the ancestral mode of reproduction in the genus, and this mating system has persisted for between 24 and 180 million years (Weeks et al. 2006b). However, the rate of selfing seems to be lower in *Eulimnadia*, and the level of inbreeding depression and maximal frequencies of males higher, than in *C. elegans* (Weeks et al. 1999; Weeks 2004). It is unknown how long *C. elegans* and *C. briggsae* have been selfing. Although the selfing species are separated by millions of years from their closest known relatives, the ecology and biodiversity of *Caenorhabditis* species is poorly understood and extensive sampling may yet identify new closely related outcrossing and self-fertile species, and narrow down the possible time since hermaphroditism evolved. Here, we have shown that natural isolates of *C. elegans* suffer from outbreeding rather than inbreeding depression. How the species arrived at this state, the time scale of this process, and whether the species will persist are intriguing problems.

ACKNOWLEDGMENTS

We are grateful to M.-A. Félix and A. Barrière for providing *C. elegans* strains. We thank M. Ailion, A. Barrière, M.-A. Félix, M. Rockman, and H. Teotónio for discussing unpublished results. We greatly appreciate the helpful comments and suggestions of D. Charlesworth, C. Haag, A. Barrière, M.-A. Félix, and two anonymous reviewers. We also thank P. Keightley, D. Halligan and A. Peters for helping with the experimental design, and I. White for assistance with the statistical analysis. This work was funded by postgraduate scholarships from the Natural Sciences and Engineering Research Council (Canada) and the University of Edinburgh School of Biological Sciences to ESD, by the Royal Society (United Kingdom) to BC, and by the National Science Foundation (U.S.A.) with International Research Fellowship Program grant 0401897 to ADC.

LITERATURE CITED

Allard, R. W. 1975. Mating system and microevolution. Genetics 79:115–126.

Armbruster, P., and D. H. Reed. 2005. Inbreeding depression in benign and stressful environments. Heredity 95:235–242.

Baird, S. E. 1999. Natural and experimental association of *Caenorhabditis remanei* with *Trachelipus rathkii* and other terrestrial isopods. Nematology 1:471–475.

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

———. 2007. Temporal dynamics and linkage disequilibrium in natural *C. elegans* populations. Genetics, *In press*.

Barton, N., and B. O. Bengtsson. 1986. The barrier to genetic exchange between hybridizing populations. Heredity 57:357–376.

Braendle, C., and M.-A. Félix. 2006. Sex determination: ways to evolve a hermaphrodite. Curr. Biol. 16:R468–R471.

Charlesworth, B. 1994. Evolution in age-structured populations. Cambridge Univ. Press, Cambridge, U.K.

Charlesworth, B., and D. Charlesworth. 1998. Some evolutionary consequences of deleterious mutations. Genetica 102/103:3–19.

Charlesworth, D. 1984. Androdioecy and the evolution of dioecy. Biol. J. Linn. Soc. 23:333–348.

———. 2003. Effects of inbreeding on the genetic diversity of populations. Philos. Trans. R. Soc. Lond. B 358:1051–1070.

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

Crnokrak, P., and S. C. H. Barrett. 2002. Purging the genetic load: a review of the experimental evidence. Evolution 56:2347–2358.

Crnokrak, P., and D. A. Roff. 1999. Inbreeding depression in the wild. Heredity 83:260–270.

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

———. 2006. Nucleotide polymorphism and linkage disequilibrium in wild populations of the partial selfer *Caenorhabditis elegans*. Genetics 172:171–184.

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

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

Demeester, L. 1993. Inbreeding and outbreeding depression in *Daphnia*. Oecologia 96:80–84.

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.

Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. Longman Group Ltd., Essex, UK.

Glémin, S., J. Ronfort, and T. Bataillon. 2003. Patterns of inbreeding depression and architecture of the load in subdivided populations. Genetics 165:2193–2212.

Goodwillie, C. 1999. Multiple origins of self-compatibility in *Linanthus* section Leptosiphon (Polemoniaceae): phylogenetic evidence from internal-transcribed-spacer sequence data. Evolution 53:1387–1395.

Goodwillie, C., S. Kalisc, and S. G. Eckert. 2005. The evolutionary enigma of mixed mating systems in plants: occurrence, theoretical explanations, and empirical evidence. Annu. Rev. Ecol. Evol. Syst. 36:47–79.

Haber, M., M. Schüngel, A. Putz, S. Müller, 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.

Halligan, D. L., A. D. Peters, and P. D. Keightley. 2003. Estimating numbers of EMS-induced mutations affecting life-history traits in *Caenorhabditis elegans* in crosses between inbred sublines. Genet. Res. 82: 191–205.

Hauser, T. P., and V. Loeschcke. 1995. Inbreeding depression in *Lychnis flos-cuculi* (Caryophyllaceae): effects of different levels of inbreeding. J. Evol. Biol. 8:589–600.

Hayes, C. N., J. A. Winsor, and A. G. Stephenson. 2005. A comparison of male and female responses to inbreeding in *Cucurbita pepo* subsp. *texana* (Cucurbitaceae). Am. J. Bot. 92:107–115.

Henry, P.-Y., R. Pradel, and P. Jarne. 2003. Environment-dependent inbreeding depression in a hermaphroditic freshwater snail. J. Evol. Biol. 16:1211–1222.

Hill, R. C., C. E. de Carvalho, J. Salogiannis, B. Schlager, D. Pilgrim, and E. S. Haag. 2006. Genetic flexibility in the convergent evolution of hermaphroditism in *Caenorhabditis* nematodes. Dev. Cell 10:531–538.

Hodgkin, J., and T. Doniach. 1997. Natural variation and copulatory plug formation in *Caenorhabditis elegans*. Genetics 146:149–164.

Holtsford, T. P., and N. C. Ellstrand. 1990. Inbreeding effects in *Clarkia tembloriensis* (Onagraceae) populations with different natural outcrossing rates. Evolution 44:2031–2046.

Jarne, P., and D. Charlesworth. 1993. The evolution of the selfing rate in functionally hermaphroditic plants and animals. Annu. Rev. Ecol. Syst. 24:441–466.

Jiang, M., J. Ruy, M. Kiraly, K. Duke, V. Reinke, and S. K. Kim. 2001. Genome-wide analysis of developmental and sex-regulated gene expression profiles in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA 98:218–223.

Johnson, T. E., and E. W. Hutchinson. 1993. Absence of strong heterosis for life-span and other life-history traits in *Caenorhabditis elegans*. Genetics 134:465–474.

Johnson, T. E., and W. B. Wood. 1982. Genetic analysis of life span in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA 79:6603–6607.

Johnston, M. O., and D. J. Schoen. 1996. Correlated evolution of self-fertilization and inbreeding depression: an experimental study of nine populations of *Amsinckia* (Boraginaceae). Evolution 50: 1478–1491.

Keightley, P. D., and A. Caballero. 1997. Genomic mutation rates for lifetime reproductive output and life span in *Caenorhabditis elegans*. Proc. Natl. Acad. Sci. USA 94:3823–3827.

Keightley, P. D., E. K. Davies, A. D. Peters, and R. G. Shaw. 2000. Properties of ethylmethane sulfonate-induced mutations affecting life-history traits in *Caenorhabditis elegans* and inferences about bivariate distributions of mutation effects. Genetics 156:143–154.

Keller, L. F., and D. M. Waller. 2002. Inbreeding effects in wild population. Trends Ecol. Evol. 17:230–241.

Kimura, M., and T. Ohta. 1971. Theoretical aspects of population genetics. Princeton Univ. Press, Princeton, NJ.

Kiontke, K., and D. H. Fitch. 2005. The phylogenetic relationships of *Caenorhabditis* and other rhabditids (August 11, 2005) *in* The *C. elegans* Research Community. WormBook, doi/10.1895/wormbook.1.11.1.

Kiontke, K., and W. Sudhaus. 2006. Ecology of *Caenorhabditis* species (January 9, 2006) *in* The *C. elegans* Research Community. Wormbook, doi/10.1895/wormbook.1.37.1

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. USA 101:9003–9008.

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

Lakowski, B., and S. Hekimi. 1996. Determination of life-span in *Caenorhabditis elegans* by four clock genes. Science 272:1010–1013.

Lande, R., and D. W. Schemske. 1985. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. Evolution 39:24–40.

Lloyd, D. G. 1975. The maintenance of gynodioecy and androdioecy in angiosperms. Genetica 45:325–339.

Lyons, E. E. 1996. Breeding system evolution in *Leavenworthia*. II. Genetic and nongenetic parental effects on reproductive success in selfing and more outcrossing populations of *Leavenworthia crassa*. Am. Nat. 147:65–85.

Nayak, S., J. Goree, and T. Schedl. 2005. *fog-2* and the evolution of self-fertile hermaphroditism in *Caenorhabditis*. PLoS Biol. 3:e6.

Otto, S. P., C. Sassaman, and M. W. Feldman. 1993. Evolution of sex determination in the crustacean shrimp *Eulimnadia texana*. Am. Nat. 141:329–337.

Pannell, J. R. 2002. The evolution and maintenance of androdioecy. Annu. Rev. Ecol. Syst. 33:397–425.

Peters, A. D., D. L. Halligan, M. C. Whitlock, and P. D. Keightley. 2003. Dominance and overdominance of mildly deleterious induced mutations for fitness traits in *Caenorhabditis elegans*. Genetics 165:589–599.

Phillips, P. C., and N. A. Johnson. 1998. The population genetics of synthetic lethals. Genetics 150:449–458.

Rieseberg, L. H., C. T. Philbrick, P. E. Pack, M. A. Hanson, and P. Fritsch. 1993. Inbreeding depression in androdioecious populations of *Datisca glomerata* (Datiscaceae). Am. J. Bot. 80:757–762.

Rocheleau, G., and S. Lessard. 2000. Stability analysis of the partial selfing selection model. J. Math. Biol. 40:541–574.

Schoen, D. J., M. O. Johnston, A.-M. L'Heureux, and J. V. Marsolais. 1997. Evolutionary history of the mating system in *Amsinckia* (Boraginaceae). Evolution 51:1090–1099.

Schulenburg, H., and S. Müller. 2004. Natural variation in the response of *Caenorhabditis elegans* towards *Bacillus thuringiensis*. Parasitology 128:433–443.

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.

Stebbins, G. L. 1957. Self-fertilization and population variability in the higher plants. Am. Nat. 91:337–354.

Stewart, A. D., and P. C. Phillips. 2002. Selection and the maintenance of androdioecy in *Caenorhabditis elegans*. Genetics 160:975–982.

Sulston, J., and J. Hodgkin. 1988. Methods. Pp. 587–606 *in* W. B. Wood, ed. The Nematode *Caenorhabditis elegans*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Takebayashi, N., and P. P. Morrell. 2001. Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. Am. J. Bot. 88:1143–1150.

Templeton, A. R. 1986. Coadaptation and outbreeding depression. Pp. 33–59 *in* M. Soulé, ed. Conservation biology: the science of scarcity and diversity. Sinauer Associates, Sunderland, MA.

Teotónio, H., D. Manoel, and P. C. Phillips. 2006. Genetic variation for outcrossing among *Caenorhabditis elegans* isolates. Evolution 60:1300–1305.

Theodorou, K., and D. Couvet. 2002. Inbreeding depression and heterosis in a structured population; influence of the mating system. Genet. Res. 80:107–116.

Truyens, S., M. M. Arbo, and J. S. Shore. 2005. Phylogenetic relationships, chromosome and breeding system evolution in *Turnera* (Turneraceae): inferences from its sequence data. Am. J. Bot. 92:1749–1758.

Uyenoyama, M. K., K. E. Holsinger, and D. M. Waller. 1993. Ecological and genetic factors directing the evolution of self-fertilization. Oxf. Surv. Evol. Biol. 9:327–381.

Vassilieva, L. L., and M. Lynch. 1999. The rate of spontaneous mutation for life-history traits in *Caenorhabditis elegans*. Genetics 151:119–129.

Vogler, D. W., K. Filmore, and A. G. Stephenson. 1999. Inbreeding depression in *Campanula rapunculoides* L. I. A comparison of inbreeding depression in plants derived from strong and weak self-incompatibility phenotypes. J. Evol. Biol. 12:483–494.

Waller, D. M. 1993. The statics and dynamics of mating system evolution. Pp. 97–117 *in* N. W. Thornhill, ed. The natural history of inbreeding and outbreeding: theoretical and empirical perspectives. Univ. of Chicago Press, Chicago.

Weeks, S. C. 2004. Levels of inbreeding depression over seven generations of selfing in the androdioecious clam shrimp, *Eulimnadia texana*. J. Evol. Biol. 17: 475–484.

Weeks, S. C., V. Marcus, and B. R. Crosser. 1999. Inbreeding depression in a self-compatible, androdioecious crustacean, *Eulimnadia texana*. Evolution 53:472–483.

Weeks, S. C., C. Benvenuto, and S. K. Reed. 2006a. When males and hermaphrodites coexist: a review of androdioecy in animals. Integr. Comp. Biol. 46:449–464.

Weeks, S. C., T. F. Sanderson, S. K. Reed, M. Zofkova, B. Knott, U. Balaraman, G. Pereira, D. M. Senyo, and W. R. Hoeh. 2006b. Ancient androdioecy in the freshwater crustacean *Eulimnadia*. Proc. R. Soc. Lond. B 273:725–734.

Whitlock, M. C. 2002. Selection, load, and inbreeding depression in a large metapopulation. Genetics 160:1191–1202.

Willis, J. H. 1999. The role of genes of large effect on inbreeding depression in *Mimulus guttatus*. Evolution 53:1678–1691.

Wilson, D. S., and M. Turelli. 1986. Stable underdominance and the evolutionary invasion of empty niches. Am. Nat. 127:835–850.

Wong, A., P. Boutis, and S. Hekimi. 1995. Mutations in the clk-1 gene of *Caenorhabditis elegans* affect developmental and behavioral timing. Genetics 139:1247–1259.

Associate Editor: D. Schoen

Heredity (2007), 1–12
© 2007 Nature Publishing Group All rights reserved 0018-067X/07 $30.00
www.nature.com/hdy

ORIGINAL ARTICLE

Hakuna Nematoda: genetic and phenotypic diversity in African isolates of *Caenorhabditis elegans* and *C. briggsae*

ES Dolgin[1], M-A Félix[2] and AD Cutter[1,3]
[1]*Institute of Evolutionary Biology, University of Edinburgh, King's Buildings, Edinburgh, UK;* [2]*Institut Jacques Monod, Centre National de la Recherche Scientifique, Universities of Paris, Paris, France and* [3]*Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada*

Caenorhabditis elegans and *C. briggsae* have many parallels in terms of morphology, life history and breeding system. Both species also share similar low levels of molecular diversity, although the global sampling of natural populations has been limited and geographically biased. In this study, we describe the first cultured isolates of *C. elegans* and *C. briggsae* from sub-Saharan Africa. We characterize these samples for patterns of nucleotide polymorphism and vulva precursor cell lineage, and conduct a series of hybrid crosses in *C. briggsae* to test for genetic incompatibilities. The distribution of genetic diversity confirms a lack of geographic structure to *C. elegans* sequences but shows genetic differentiation of *C. briggsae* into three distinct clades that may correspond to three latitudinal ranges. Despite low levels of molecular diversity, we find considerable variation in cell division frequency in African *C. elegans* for the P3.p vulva precursor cell, and in African *C. briggsae* for P4.p, a variation that was not previously observed in this species. Hybrid crosses did not reveal major incompatibilities between *C. briggsae* strains from Africa and elsewhere, and there was some evidence of inbreeding depression. These new African isolates suggest that important ecological factors may be shaping the patterns of diversity in *C. briggsae*, and that despite many similarities between *C. elegans* and *C. briggsae*, there may be more subtle differences in their natural histories than previously appreciated.
Heredity advance online publication, 12 December 2007; doi:10.1038/sj.hdy.6801079

Keywords: *Caenorhabditis elegans*; *Caenorhabditis briggsae*; genetic variation; vulva; self-fertilization; androdioecy

Introduction

As a model organism, the nematode *Caenorhabditis elegans* has provided an outstanding system for elucidating fundamental problems in biology. However, the attitude regarding the natural history and ecology of *Caenorhabditis* has long been 'hakuna matata', to use the Swahili phrase meaning 'no worries'. To achieve a robust understanding of the evolutionary context in which *C. elegans*' developmental and genetic patterns emerged, it is necessary to characterize representative natural populations on a worldwide scale and to conduct comparative analyses of *Caenorhabditis* species. Specifically, the quantification of genetic and phenotypic variation in species with shared common ancestry permits inference about the evolutionary forces and population processes that shape the life history and development of these organisms.

Along with *C. elegans*, *C. briggsae* forms part of a monophyletic clade of five species in laboratory culture known as the *Elegans* group of the genus *Caenorhabditis* (Kiontke and Fitch, 2005). Like *C. elegans*, *C. briggsae* is androdioecious (self-fertile hermaphrodites and facultative males), but both species evolved independently from gonochoristic (male–female) ancestors (Kiontke *et al*., 2004). Despite analogous breeding systems and similar morphology (Nigon and Dougherty, 1949), their genome sequences have diverged drastically (Coghlan and Wolfe, 2002; Stein *et al*., 2003), although they do show a high degree of chromosomal synteny (Hillier *et al*., 2007). The development of *C. briggsae* as a laboratory system is still in its infancy, but with a growing number of genetic resources becoming available, including a sequenced genome and recombination maps, this species is emerging as a useful companion species of *C. elegans* (Baird and Chamberlin, 2006).

C. elegans and *C. briggsae* are cosmopolitan species with overlapping geographic distributions, with at least one of the two species found on each continent where intensive field collections have been carried out. Both species show similarly low levels of genetic variation, especially when compared to the related gonochoristic species, *C. remanei* (Graustein *et al*., 2002; Jovelin *et al*., 2003; Haag and Ackerman, 2005; Cutter, 2006; Cutter *et al*., 2006a, b). The predominantly self-fertilizing breeding systems shared by both *C. elegans* and *C. briggsae* has probably driven this decrease in diversity as a consequence of reduced effective population size, increased homozygosity and greater population subdivision (Charlesworth, 2003). Despite their similarities, however, *C. elegans* and

Correspondence: ES Dolgin, Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, King's Buildings, Edinburgh, EH9 3JT, UK.
E-mail: elie.dolgin@ed.ac.uk
Received 29 June 2007; revised 8 October 2007; accepted 8 November 2007

C. briggsae probably occupy different ecological niches. For example, one difference relevant to ecology is that *C. briggsae* tolerates and proliferates at higher temperatures than *C. elegans* (Fodor *et al*., 1983). *C. elegans* is also strongly induced to form the dormant dauer larval stage at high temperatures, whereas *C. briggsae* appears to lack this response (Inoue *et al*., 2007).

Another difference of potential ecological significance is the relative importance of geography in structuring genetic variation. Phylogenetic reconstructions of molecular polymorphisms among available global samples show no strong signature of geographic structure in *C. elegans* (Denver *et al*., 2003; Barrière and Félix, 2005; Haber *et al*., 2005; Cutter, 2006), but a clear division is observed in *C. briggsae* between strains from temperate latitudes and strains found near the Tropic of Cancer (Graustein *et al*., 2002; Cutter *et al*., 2006b). A major problem with our current understanding of natural variation and biogeography, however, is that it is unclear whether the focal samples adequately reflect global diversity, because sampling efforts have not targeted most regions outside North America and Europe. Although the genus was first described from specimens isolated in North Africa (Maupas, 1900), these strains were never cultured, and, to date, no *C. elegans* or *C. briggsae* strains have been analyzed from Africa or South America. To address this issue, one of us (ESD) undertook intensive field collections in two countries in sub-Saharan Africa: Kenya and South Africa.

In this study, we test the hypotheses concerning geographic structuring of genotypes in *C. elegans* and *C. briggsae* with new samples from Africa. We also examine vulval cell lineages for phenotypic polymorphisms and conduct intraspecific crosses in *C. briggsae* to test for genetic incompatibilities. Our results confirm the lack of global geographic structure in the sequences of *C. elegans*, but provide further evidence for latitudinal clades in *C. briggsae*. We also find unique patterns of vulval cell division in South African strains of *C. briggsae* that have not been observed previously in this species, while crosses between isolates from disparate localities fail to demonstrate hybrid incompatibilities (Dolgin *et al*., 2007), and instead show evidence of inbreeding depression. We discuss how these new isolates add to our understanding of global variation and how evolutionary processes might act differently on the two species after their independent transitions to androdioecy.

Methods

Nematode populations

We isolated 38 new strains of *C. elegans* and 25 new strains of *C. briggsae* from single wild hermaphrodites in Kenya and South Africa in March–April 2006. Despite extensive sampling from 36 sites, *Caenorhabditis* nematodes were only found in four locations: (1) compost from a private garden (Johannesburg, South Africa); (2) compost from a plant nursery (Ceres, South Africa); (3) compost from a mushroom farm (Limuru, Kenya) and (4) leaf litter in a public park (Nairobi, Kenya). Strain name designations for *C. elegans* are ED3040–ED3077, and for *C. briggsae* are ED3078–ED3102. The sampling sites and strain designations are described in Table 1. Various other nematode species were also found in these sites, including many strains of *Oscheius tipulae*, which have also been characterized (Baïlle *et al*., 2008).

The sampling protocol was carried out using the 'Isolation on an Agar Plate' technique as described by Barrière and Félix (2006). Briefly, small samples of compost or leaf litter were placed on 9-cm NGM-lite agar plates spotted with *Escherichia coli* OP50. These plates were then monitored for nematodes crawling out into the bacterial lawn for 24–48 h, and individual worms were isolated onto separate 6-cm plates to establish independent iso-hermaphrodite lineages. Self-fertile individuals were examined for morphological features under compound microscopy, and species identity was made through mating tests and by sequencing a portion of the small subunit ribosomal RNA gene (Floyd *et al*., 2002).

For DNA sequencing, we randomly chose a subset of 16 strains of each species from the African collection as representative of these sampling sites (Table 1). The sequences from these strains were then compared to isolates from a larger worldwide data set of 118 *C. elegans* strains (Cutter, 2006) and 63 *C. briggsae* strains (Cutter *et al*., 2006b). To gain a more complete understanding of the molecular diversity in the two species, we also used an additional nine strains that were recently isolated from geographically unique locations and had not previously been assessed for nucleotide polymorphism: a *C. elegans* strain from Madagascar (LKC34; isolated by V. Stowell and donated by L. Carta), a *C. elegans* strain from Ribeiro Frio, Madeira (JU258; isolated by MAF), a *C. elegans* strain from Kakegawa, Japan (JU1088; isolated by MAF), two *C. elegans* strains from Concepcion, Chile (JU1171 and JU1172; isolated by MAF), a *C. briggsae* strain from Kruger National Park, South Africa (DF5100; isolated by W Sudhaus and donated by K Kiontke), a *C. briggsae* strain from Kakegawa, Japan (JU1085; isolated by MAF) and two *C. briggsae* strains from Salt Lake City, UT, USA (EG4181 and EG4207; isolated by M Ailion and obtained from the Caenorhabditis Genetics Center). The geographic origin of each *C. elegans* and *C. briggsae* strain included in this analysis is shown in Figure 1.

Molecular methods

DNA from single individuals was obtained using a NaOH digestion protocol (Floyd *et al*., 2002). We sequenced the gene fragments that had previously been analyzed for a larger global collection of strains (Cutter, 2006; Cutter *et al*., 2006b). These comprised six genes from *C. elegans* chromosomes II and X, and the putative orthologs of five of these genes plus one locus nearby the sixth gene in *C. briggsae*, with primers designed to span a long (>500 bp) intron. Names, lengths, positions and primers of the sequenced loci can be found as described by Cutter (2006) and Cutter *et al*. (2006b). Both strands were sequenced on an ABI Prism 3730 automated sequencer at the School of Biological Sciences Sequencing Service, University of Edinburgh.

Sequence analysis

Sequence alignment and manual editing to remove the primers were performed with Sequencher 4.6 and BioEdit 7.0.5. Calculations of diversity (from pairwise differences, π, and from the number of segregating sites, θ), Tajima's *D* tests of neutrality (Tajima, 1989), and tests

Table 1 New *C. elegans* and *C. briggsae* strains isolated in Kenya and South Africa, and a list of the strains used for DNA sequencing

Location (nearest city)	*Latitude (S)*	*Longitude (E)*	*Altitude (m)*	*C. elegans*			*C. briggsae*		
				No.	*Strain designations*	*Strains used in sequencing (and haplotypes)*	*No.*	*Strain designations*	*Strains used in sequencing*
Johannesburg, South Africa	26°10′	28°01′	1669	1	ED3040	ED3040 (α)	12	ED3078–ED3089	ED3078 ED3079 ED3080 ED3082 ED3083 ED3084 ED3087 ED3089
Ceres, South Africa	33°22′	19°19′	478	12	ED3041–ED3052	ED3042 (β) ED3046 (γ) ED3048 (β) ED3049 (γ) ED3051 (β) ED3052 (δ)	0		
Limuru, Kenya	1°05′	36°39′	2356	24	ED3053–ED3076	ED3054 (ε) ED3057 (ε) ED3063 (ε) ED3066 (ε) ED3072 (ε) ED3073 (ε) ED3075 (ε) ED3076 (ε)	0		
Nairobi, Kenya	1°19′	36°48′	1708	1	ED3077	ED3077 (β)	13	ED3090–ED3102	ED3091 ED3092 ED3095 ED3096 ED3098 ED3099 ED3100 ED3101

C. elegans haplotype designations are as shown in Figure 2a.

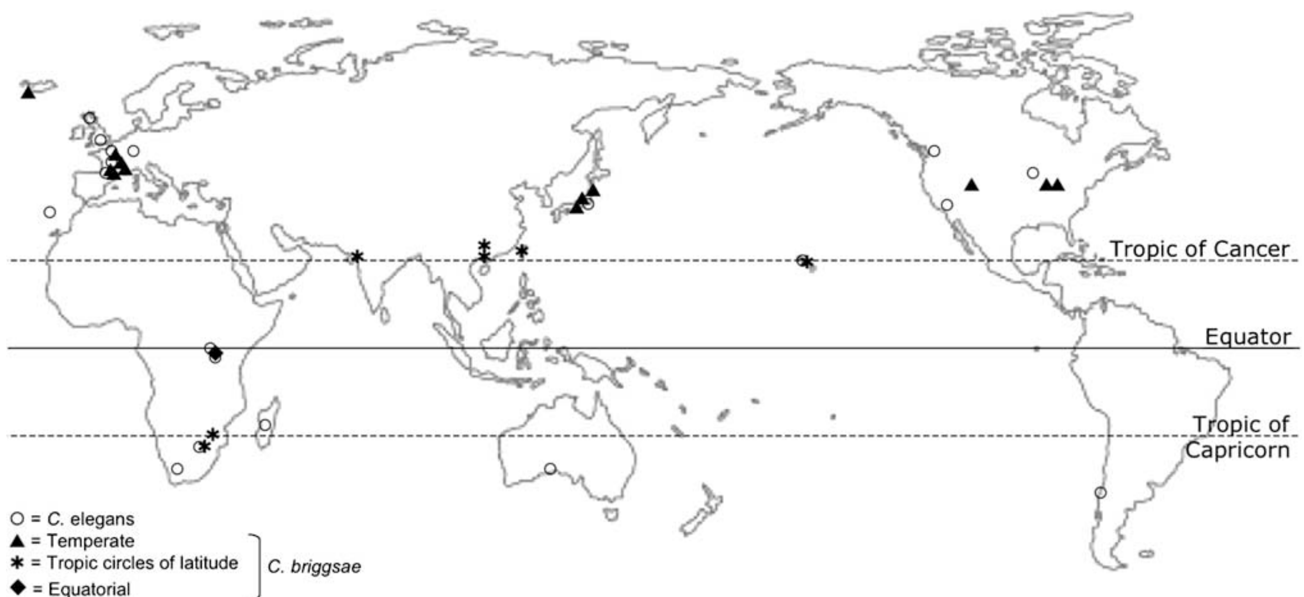


Figure 1 Geographic distribution of strains used for molecular analysis. Open circles indicate *C. elegans* strains and closed symbols indicate *C. briggsae* strains from different latitudinal clades: triangles (temperate latitude samples), asterisks (Tropic circles of latitude samples) and diamonds (equatorial samples). Note that the exact sampling location of *C. elegans* strain LKC34 is uncertain, although it is known to originate from Madagascar.

of population differentiation using average values of K^*_{st} among loci (Hudson *et al*., 1992) and Hudson's (Hudson, 2000) nearest-neighbor statistic, S_{nn}, from concatenated sequences were made using DnaSP 4.10.9 (Rozas *et al*., 2003). We present diversity data from silent sites (synonymous and intronic positions); sites corresponding to indels or incomplete data were excluded from the analyses. Neighbor networks and neighbor-joining trees were constructed with concatenated sequences using SplitsTree 4.6 (Huson and Bryant, 2006).

Vulval cell lineage

The frequency of division of the vulval precursor cells, P3.p and P4.p, was measured in a subset of the new African isolates—five *C. elegans* strains (ED3040, ED3046, ED3052, ED3054 and ED3077) and four *C. briggsae* strains (ED3082, ED3087, ED3092 and ED3101)—in addition to *C. elegans* strains JU258 and LKC34. The results were compared to a global collection of 13 *C. elegans* strains and 6 *C. briggsae* strains that had previously been scored for these polymorphic phenotypes (Delattre and Félix, 2001). For some of these previously analyzed strains, more worms were also scored to obtain larger sample sizes. We also screened eight strains from other species in the *Elegans* group of the *Caenorhabditis* genus for comparison—four strains of *C. remanei* (JU724, JU825, PB4641 and SB146), two strains of *C. brenneri* (CB5161 and PB2801) and two strains of *Caenorhabditis* sp. 5 (JU727 and SB378). The fate of the Pn.p cells was assessed at the L4 stage using Nomarski miscroscopy under standard conditions. We tested for heterogeneity between intraspecific strains by using replicated goodness-of-fit tests on the number of individuals in which the Pn.p cell divided or not, and performed *post hoc* pairwise tests after applying the Dunn–Šidák correction (Sokal and Rohlf, 1995, p. 239).

Hybrid crosses

To test for possible incompatibilities between strains of *C. briggsae*, a series of crosses were set up using strains from South Africa (ED3083) and Kenya (ED3101), as well as three additional *C. briggsae* strains: AF16 from India, HK104 from Japan and ED3034 from Taiwan. Pairwise crosses to the two African strains created seven crosses: ED3083 × ED3034, ED3083 × AF16, ED3083 × HK104, ED3083 × ED3101, ED3101 × ED3034, ED3101 × AF16 and ED3101 × HK104. We also performed an eighth cross between AF16 and HK104 since this interstrain cross has been studied previously (Baird *et al*., 2005). Hybrid crosses were only performed in *C. briggsae* in this study because only African samples of this species showed markedly different molecular and developmental patterns (see the Results), and hybrid crosses with representative populations have been performed elsewhere for *C. elegans* (Dolgin *et al*., 2007).

In each cross, we generated mixed-mating populations of males (M) and hermaphrodites (H) for each strain. We then established pure-strain F1s from each strain (for example, ED3101 × ED3101 and AF16 × AF16), and the two reciprocal F1 hybrids (for example, ED3101–M × AF16–H and AF16–M × ED3101–H). We then measured the number of surviving late larval F2s (to give a measure of F1 brood size), the percentage of F2 embryonic lethality and the developmental timing to reach the L4 larval stage in the F2s. Interstrain crosses of *C. briggsae* were performed using the methods that Dolgin *et al*. (2007) used for *C. elegans*, with a few notable exceptions. Following synchronization of the pure-strain parental worms using alkaline hypochlorite, some L4 hermaphrodites were transferred to a new plate and the remaining worms were left with the males in a mixed-sex population. The isolated hermaphrodites were permitted to self for ~3.5 days following first egg-lay to exhaust them of self-sperm. Next generation males were then taken from the mixed-sex population and were mated to these sperm-depleted hermaphrodites for a 24 h interval. The mated hermaphrodites were subsequently transferred to new plates to lay eggs for 5 h. From these F1 progeny, many L4 hermaphrodites of each pure strain or hybrid genotype (mean number = 19.2 ± 1.1 s.e.) were set up individually to measure surviving larval brood size and embryonic lethality. These worms were transferred every 24 h to a new plate; 36 h after transfer, unhatched eggs and surviving larvae were counted. Subsets of the remaining F1 hermaphrodites were transferred *en masse* to new plates at the L4 stage. Several hours after eggs were observed on these plates, the worms were transferred twice again to new plates to lay eggs for two 1 h intervals. These plates were then monitored for when the F2 progeny molted from the L4 stage to adulthood, as a measure of developmental timing. The mean brood sizes and levels of embryonic lethality within the same hybrid cross were compared using *t*-tests. Differences in developmental timing were evaluated using goodness-of-fit tests on the number of 'normal' or 'late' molting larvae, where late was defined as more than 5 h after the peak molting time.

Results

Molecular polymorphism

The African sample of *C. elegans* shows similar low levels of within-locality variation as reported previously for other regions of the world. Overall silent-site nucleotide polymorphism in the African localities was equivalent to the European samples, with Africa-wide π_{si} and θ_{si} both estimated to be ~0.2% (Table 2). Considering diversity levels of South African or Kenyan isolates independently also resulted in quantitatively similar diversity levels to that found in individual European localities. Despite low levels of variation, there is no evidence of non-neutral demographic or selective processes in the African samples, either as a whole or by considering each country individually, as there was a range of positive and negative values of Tajima's *D* across the different loci, with no values significantly different from zero (results not shown).

The 17 African samples of *C. elegans* (including LKC34 from Madagascar) contained six different haplotypes. Although these haplotypes were endemic to African samples, most polymorphic sites are also found elsewhere in the world (Figure 2a). All the African strains were generally similar to other strains and showed no strong patterns of geographic structure (Figure 3a and Supplementary Figure S1a); however, we did find some evidence of population differentiation between European and African samples ($K^*_{st} = 0.075$; $S_{nn} = 0.97$, $P < 0.0001$) and between Kenyan and South African samples

Table 2 Summary of nucleotide polymorphisms for African *C. elegans* and *C. briggsae* isolates

Locus (chromosome)	*African population samples*							*South Africa*		*Kenya*	
	Silent sites	*S*	*h*	H_d	*Indels*	π_{si}	θ_{si}	π_{si}	θ_{si}	π_{si}	θ_{si}
C. elegans											
Y25C1A.5 (II)	658.8	7	4	0.596	4	0.0036	0.0031	0.0033	0.0031	0.0013	0.0022
ZK430.1 (II)	542.8	1	2	0.529	1	0.0010	0.0005	0	0	0.0004	0.0007
E01G4.6 (II)	521.4	16	3	0.324	2	0.0066	0.0091	0.0092	0.0125	0	0
D1005.1 (X)	480.0	0	1	0	0	0	0	0	0	0	0
R160.7 (X)	605.2	0	1	0	0	0	0	0	0	0	0
T24D11.1 (X)	651.0	1	2	0.309	0	0.0005	0.0005	0.0007	0.0006	0	0
Concatenated	3459.2	25	6	0.743	7	0.0019	0.0021	0.0021	0.0026	0.0003	0.0005
Average	576.5	4.2	2.2	0.293	1.2	0.0019	0.0022	0.0022	0.0027	0.0003	0.0005
C. briggsae											
p09	484.7							0	0	0	0
p10	581.3							0.0004	0.0006	0.0004	0.0007
p11	571.3							0	0	0	0
p12	623.7							0.0007	0.0012	0	0
p13	593.5							0	0	0	0
p14	599.5							0	0	0	0
Concatenated	3452.0							0.0002	0.0003	0.0001	0.0001
Average	575.7							0.0002	0.0003	0.0001	0.0001

Abbreviations: S, segregating sites; h, haplotypes; H_d, haplotype diversity; π_{si}, silent-site diversity from pairwise differences; θ_{si}, silent-site diversity from the number of segregating sites.
Diversity estimates were calculated for each country in both species, and for all African samples for *C. elegans* but not for *C. briggsae* due to the distinct geographical structuring between countries in this species. For *C. elegans*, there were a total of 17 strains—7 from South Africa, 9 from Kenya and 1 from Madagascar (LKC34). For *C. briggsae*, there were a total of 17 strains—9 from South Africa (including DF5100) and 8 from Kenya.

($K^*_{st} = 0.215$; $S_{nn} = 0.85$, $P = 0.003$). Nonetheless, after including the African samples, we did not find evidence of any additional recombination events to those found previously, as indicated by the four-gamete test (Hudson and Kaplan, 1985). We also sequenced loci for the first known *C. elegans* isolates from Asia (JU1088 from Japan) and South America (JU1171 and JU1172 from Chile), plus the Madeiran strain JU258, since it was previously reported that this strain was highly differentiated from other strains (Haber *et al*., 2005; Stewart *et al*., 2005). Despite the unique geographical origin of these additional strains, we failed to find high divergence of any of them from global samples for these six loci (see haplotype Y for JU1088, haplotype I for JU1171, haplotype K for JU1172 and haplotype Z for JU258, in Figures 2a and 3a). Furthermore, JU258 was not strongly allied to the Hawaiian strain, CB4856 (haplotype T), relative to the N2 strain (haplotype A), as previously suggested (Haber *et al*., 2005; Stewart *et al*., 2005; Maydan *et al*., 2007).

The African *C. briggsae* samples show a markedly different pattern from that found for *C. elegans*. We observed very little within-locality diversity across the six loci tested, with South African and Kenyan π_{si} and θ_{si} estimated to be extremely low (Table 2). This results from the fact that seven of eight Kenyan strains were identical, with only a single nucleotide polymorphisms (SNP) distinguishing ED3092, and of the nine South African strains, the eight strains from Johannesburg were identical to each other, and DF5100, from Kruger National Park, had only three SNP differences from the Johannesburg strains. Nonetheless, the African *C. briggsae* haplotypes were very informative about geographic structure (Figure 2b). The South African haplotypes were remarkably similar to those of other strains found in the northern hemisphere near the Tropic of Cancer, and the Kenyan haplotypes were unlike those of any other strains previously described, with many newly identified polymorphisms not shared with other populations (Figure 3b and Supplementary Figure S1b). For example, the p11 locus had previously been shown to be monomorphic (Cutter *et al*., 2006b), but we found seven SNPs and three indels in the Kenyan samples (Figure 2b). However, since most of these polymorphisms were unique to this locality, we found no further evidence of historical recombination after including the African samples, as indicated by the four-gamete test (Hudson and Kaplan, 1985).

The 'tropical' strains previously described originated from China, India and Hawaii—all situated within 3° of the Tropic of Cancer. The South African strains described here were isolated in Johannesburg, less than 3° south of the Tropic of Capricorn, and Kruger National Park, which straddles the Tropic of Capricorn. Considering that these samples were all quite different from the Kenyan samples, which is also technically a tropical location, we propose that what was previously called 'tropical' strains (Cutter *et al*., 2006b) should be considered as 'Tropic circles of latitude' strains, and suggest that molecular diversity in *C. briggsae* strains may be partitioned into three major latitudinal clades: temperate samples, Tropic circles of latitude samples and equatorial

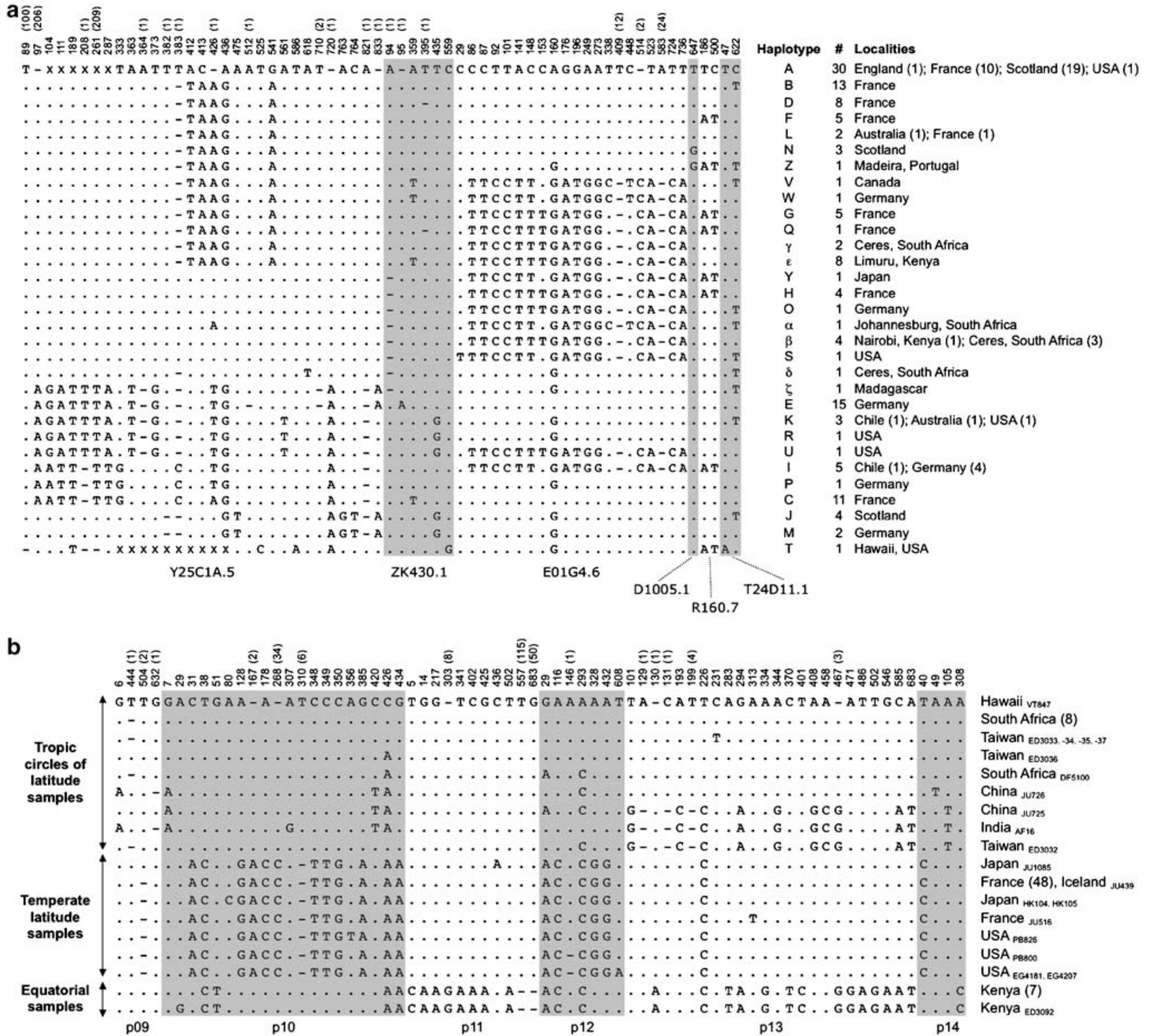


Figure 2 Summary of multilocus haplotypes for *C. elegans* (**a**) and *C. briggsae* (**b**). Loci are labeled along the bottom, with the positions of polymorphisms relative to the start of each locus indicated across the top (indel lengths in parentheses). For *C. elegans*, the haplotype designations follow those described by Cutter (2006), with JU1088 denoted by haplotype Y, JU258 denoted by haplotype Z and the African haplotypes indicated by Greek letters. For a list of which strains correspond to each African haplotype, see Table 1. Haplotype X described by Cutter (2006) was found to be equivalent to A, as a consequence of an alignment error at position 426 in Y25C1A.5. We also amend several presentation errors from Figure 1 as described by Cutter (2006).

samples (Figures 2b and 3b). An alternative model of geographic structure is that a variety of deeply divergent clades of *C. briggsae* populate the globe in a manner that only mimics a latitudinal distribution, as a consequence of available sampling. We also sequenced three additional recently isolated strains: a strain from Kakegawa, Shizuoka Prefecture, Japan (JU1085) and two from Salt Lake City, UT, USA (EG4181 and EG4207), and observed that these strains were quite similar to other strains from temperate latitudes, consistent with the prediction of a temperate latitudinal clade (see Figures 2b and 3b). More extensive global sampling is required to confirm the generality of this latitudinal geographic structure to genetic diversity in *C. briggsae*.

Vulval cell lineage

In *C. elegans*, the significant variation we observed among strains in the division frequency of the P3.p vulval precursor cell ($\chi^2_{19} = 447.5$, $P < 0.0001$) indicates that considerable heritable phenotypic diversity exists in the African sample, despite a lack of substantial genetic variation at the loci surveyed here (Figure 4). For example, ED3046 (South Africa) and ED3054 (Kenya) exhibited a 38% difference in P3.p division frequency, yet differed by only one SNP across the ~4-kb molecular dataset (see Figure 2a). Unlike the variation known for P3.p cell fate, the P4.p cell was previously thought to be largely invariant in the *Caenorhabditis* genus (Delattre and Félix, 2001); however, we found significant variation in

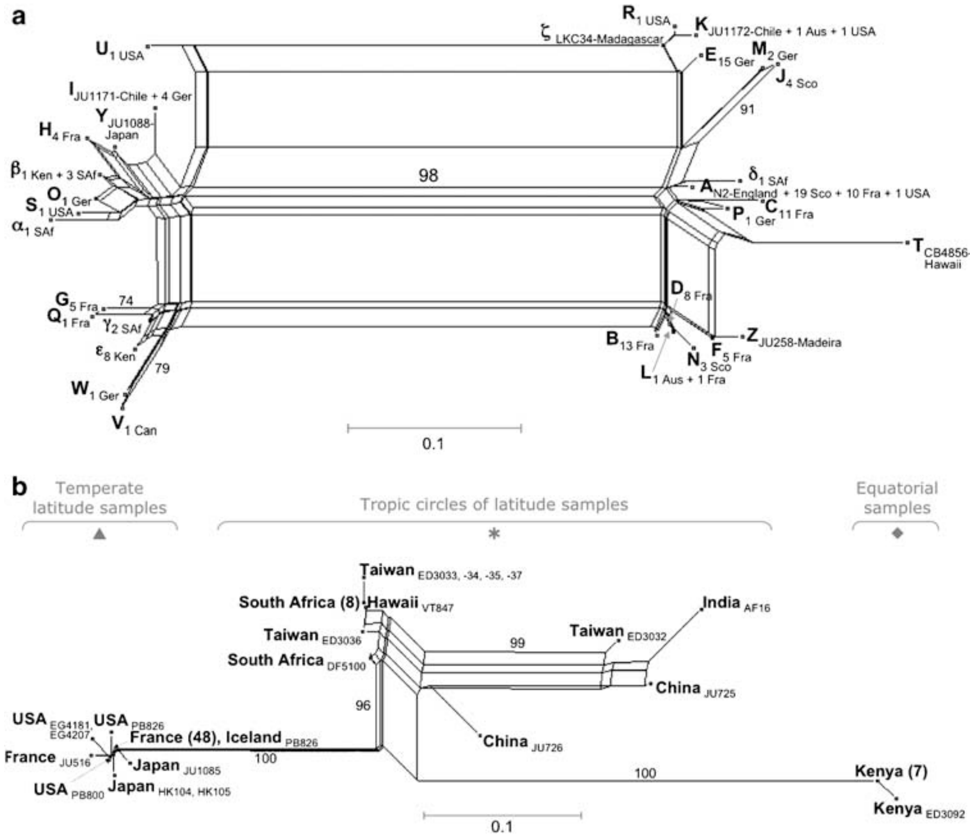


Figure 3 Unrooted *p*-distance neighbor-nets for *C. elegans* (**a**) and *C. briggsae* (**b**) based on concatenated multilocus haplotypes. *C. elegans* haplotype designations are as shown in Figure 2a; African haplotypes are designated by Greek letters; subscripts indicate the number of strains per haplotype and the country of origin (Aus, Australia; Can, Canada; Fra, France; Ger, Germany; Ken, Kenya; SAf, South Africa; Sco, Scotland). *C. briggsae* clades are labeled following the latitudinal clades with the symbols in parentheses corresponding to those used to represent the strains in Figure 1. Bootstrap values ⩾70% of 1000 replicates are shown next to the branches. Neighbor-joining trees are shown in Supplementary Figure S1.

C. briggsae strains for both the P3.p ($\chi^2_9 = 33.0$, $P = 0.0001$) and P4.p cells ($\chi^2_9 = 94.2$, $P < 0.0001$). In the two South African *C. briggsae* strains tested (ED3082 and ED3087), P4.p failed to divide 23–24% of the time. This level was significantly greater than for any other *C. briggsae* strain (all pairwise tests, $P < 10^{-8}$). In our screens of three other species in the *Elegans* group of the genus, we observed this level of P4.p nondivision only in *C. briggsae*'s closest relative, *C.* sp. 5 (JU727, 67% P4.p nondivision frequency, $n = 69$; SB378, 5% P4.p nondivision frequency, $n = 113$). The unusual P4.p cell-lineage features found in the South African *C. briggsae* strains highlights the potential for significant phenotypic differences still to be found in these isolates, despite a high sequence similarity with the northern tropical circles of latitude strains.

Hybrid crosses

Despite large differences in the molecular sequences of *C. briggsae* isolates from different parts of the world, we did not find any gross hybrid incompatibilities between African strains and strains from elsewhere, although there is substantial quantitative variation between strains for brood size and embryonic lethality (Figure 5). Comparisons of pure strain versus hybrid F1s, however, show some evidence of mid-parent heterosis (MPH) for brood size, in which the hybrids show increased levels over the mean of the two parents. MPH was significantly positive in three of eight crosses, and the mean MPH value across all the crosses was 16.7%. The F1 hybrids of one cross, ED3083 (South Africa) × ED3034 (Taiwan), even had significantly greater brood sizes than either pure-strain parent, displaying best-parent heterosis ($t_{30} = 2.07$, $P = 0.047$), although this appears to be due to a large reduction in the embryonic lethality in the hybrids, which is quite high for both pure strains (Figure 5).

We repeated the assays of brood size and embryonic lethality in pure-strain worms, and confirmed that these same strains displayed high proportions of dead embryos and reduced fecundities (results not shown). The low brood sizes in these strains was also independently observed elsewhere (D Denver, personal communication). In these strains (and others from Johannesburg, South Africa), we also observed that the worms suffered from pale and vacuolated intestines (Supplementary Figure S2). This phenotype could result from an intracellular pathogen of intestinal cells, yet it was resistant to bleaching with alkaline hypochlorite—a method that is known to kill many extracellular pathogens, and eliminated an intracellular bacterium previously described in *C. elegans* (Barrière and Félix, 2005).

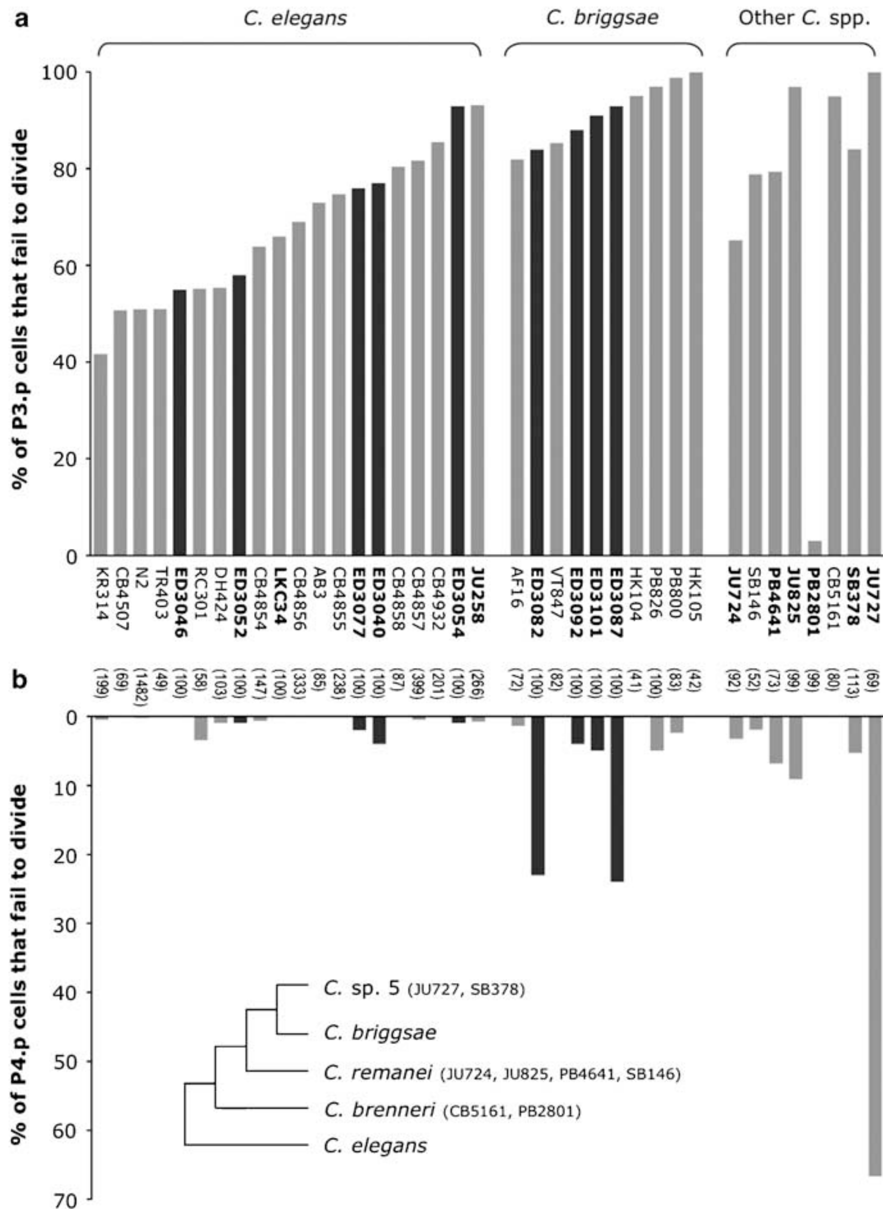


Figure 4 Percentage of P3.p (**a**) and P4.p (**b**) nondivision in hermaphrodites of *C. elegans*, *C. briggsae* and three other *Caenorhabditis* species of the *Elegans* group of the genus. The strain names in bold denote strains not previously analyzed by Delattre and Félix (2001). The darker shaded bars highlight the results for the newly collected African isolates. The number of observed worms is shown next to the strain name. The phylogenetic topology of *Caenorhabditis* species follows Kiontke and Fitch (2005) with *C.* sp. 5 added according to K Kiontke and D Fitch (personal communication).

This raises the possibility of a genetic basis to these traits, and more work is needed to explain the abnormal intestines and poor performance of these worms.

We also monitored the duration of development from egg to L4–adult molt in the pure strain and hybrid F2s, but only observed large numbers of delayed F2s in one class of worms—the hybrids of AF16 (India) and HK104 (Japan). There was a significant difference between the duration of development of AF16 × HK104 hybrids and pure strains ($\chi^2_1 = 33.7$, $P < 0.0001$), with $\sim$20% of hybrid F2s exhibiting delayed development. This confirms previous observations (S Baird, unpublished data), and suggests that the effect is peculiar to these genotypes, since this phenomenon was not found in any other crosses involving either the African isolates (results not shown) or other crosses without AF16 (S Baird, personal communication).

Discussion

The discovery of *C. elegans* and *C. briggsae* in sub-Saharan Africa, as well as additional strains from around the world, reinforces the notion that these are indeed cosmopolitan species and offers some novel insights about levels and patterns of global variation. In *C. elegans*, this study confirms the low levels of overall silent-site polymorphism previously found in Europe for another part of the world (Cutter, 2006). In *C. briggsae*,

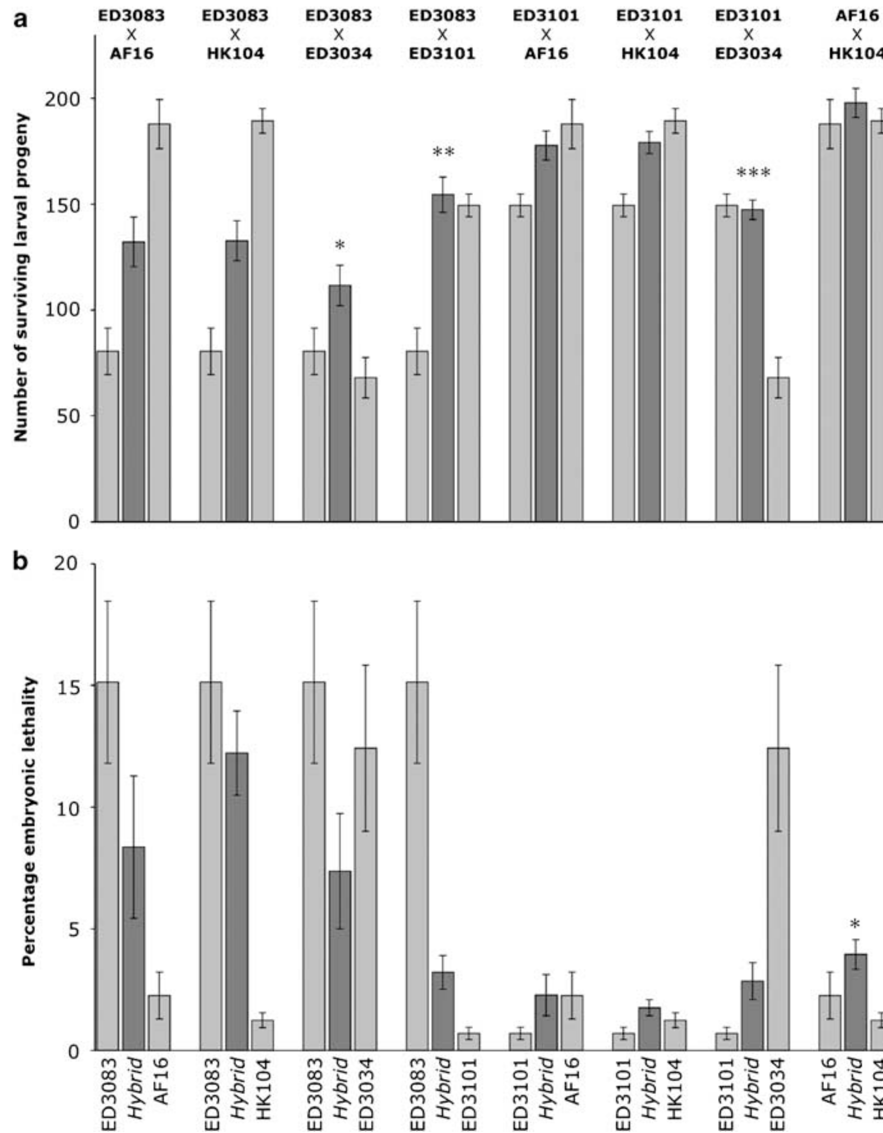


Figure 5 Surviving larval brood sizes (**a**) and embryonic lethalities (**b**) in the *C. briggsae* hybrid crosses. Pure strains (lightly shaded bars) are shown next to the hybrids (darkly shaded bars) in each cross. For ease of comparison, identical data for pure strains is shown multiple times. Asterisks (*) indicate significance in a *t*-test of the deviation between hybrid and mid-parent. *$P<0.05$, **$P<0.01$, ***$P<0.001$. Error bars indicate ±1 s.e.

Kenyan samples raise global species diversity estimates as a consequence of between population differences, despite a lack of variation within each of the two sampling sites in Africa. This result is similar to the only other known study that has established multiple wild strains from single locations (Cutter *et al*., 2006b), although whether this reflects true levels of within-population variation in *C. briggsae* or a biased local sampling procedure remains to be seen. The natural habitat of *C. elegans* and *C. briggsae* is unknown, although both species are routinely found in anthropogenic habitats, such as compost and garden soil, and they often form phoretic and/or necromenic associations with other invertebrates (Kiontke and Sudhaus, 2006). Given their apparent similarity in sampled habitats, it is striking that the two species have such different geographic patterning of molecular diversity (see Figure 3), even more so considering that *C. elegans* and *C. briggsae* were found co-occurring in the same sampling sites in Kenya and South Africa, as well as France (Barrière and Félix, 2005) and Japan (JU1085 and JU1088, described in this study). This begs the question of what could have led to the different biogeographic structures in the two species.

Population genetic studies in *C. elegans* have demonstrated that migration occurs over surprisingly large distances, including continents (Haber *et al*., 2005; Cutter, 2006; Barrière and Félix, 2005, 2007). Since the latitudinal clades we observe appear circumnavigatory, this may be true of *C. briggsae* as well. It has been proposed that dispersion in these species might be facilitated by human activities (Kiontke and Sudhaus, 2006), but it is conceivable that the migration patterns and the role of human-mediated dispersion in the two species are quite

different. Although both species co-occur in some localities, there may be substantial differences in their natural ecology that we have failed to appreciate. Another possibility is that *C. briggsae* more recently colonized different parts of the world than *C. elegans* (Cutter *et al*., 2006b). However, our evidence of a third distinct geographic haplotype group in equatorial Kenya with many unique polymorphisms makes it difficult to determine the source of such a founder event. Furthermore, the lack of additional recombination events detected from the inclusion of Kenyan isolates implies that these populations might be isolated relative to other known samples. In *C. elegans*, two strains in particular (CB4856/Hawaii and JU258/Madeira), both from remote island locations, have been shown to be highly divergent from most other strains (Koch *et al*., 2000; Haber *et al*., 2005; Stewart *et al*., 2005; although across the six loci tested here, we did not observe this for JU258, see Figures 2a and 3a), but it seems unlikely that samples from urbanized areas of Africa will be as isolated as these island strains. The absence of detectable recombination between the three clades in *C. briggsae* suggests that migration patterns and outcrossing rates could indeed be quite different in *C. briggsae* and *C. elegans*. This is further indicated by the lack of within-locality diversity in Kenya, South Africa and France (Cutter *et al*., 2006b), which implies that gene flow is not sufficiently prevalent to prevent structuring by genetic drift and/or local selection, and is qualitatively consistent with a lower rate of outcrossing in *C. briggsae* (Cutter *et al*., 2006b).

Our evidence for three distinct latitudinal haplotype groups in *C. briggsae*, including a 'Tropic circles of latitude' clade that spans both hemispheres, suggests that ecological factors might also play a role in driving the observed patterns of biogeographic structure. One might speculate that this geographical partitioning of haplotypes reflects local adaptation. One possible outcome of this scenario would be outbreeding depression. However, we observed marginal inbreeding depression in the hybrid crosses between different latitudinal clades. In fact, outbreeding depression was previously observed for *C. elegans*, although this was probably due to the effects of the hybrid breakdown of coadapted gene complexes, rather than to effects of local adaptation (Dolgin *et al*., 2007). Thus, it may be the case that local inbreeding and genetic drift are sufficient to explain the inbreeding depression observed between disparate population samples of *C. briggsae*. A possible methodological explanation for the inbreeding depression is that we performed our hybrid crosses under standard laboratory conditions of 20 °C, and inbreeding depression was primarily found for crosses involving ED3101 from Nairobi, Kenya. Considering that the average monthly maximum ambient temperature in Nairobi is never this low, these conditions might constitute a stressful environment for this strain, resulting in inbreeding depression (Armbruster and Reed, 2005), whereas experimentation under more benign conditions could produce different results.

Even though *C. elegans* displays low levels of molecular variation, wild isolates show phenotypic variation in many potentially ecologically relevant traits, including dauer sensitivity (Viney *et al*., 2003), response to pathogens (Schulenburg and Müller, 2004), clumping behavior (Hodgkin and Doniach, 1997; de Bono and Bargmann, 1998) and locomotory speed (de Bono and Bargmann, 1998). Relating these phenotypes to their ecological importance, however, remains elusive. The low degree of variation and extensive linkage disequilibrium found in *C. elegans* make it difficult to detect signatures of local adaptation using traditional molecular population genetics approaches. The same is true for *C. briggsae*, but the more structured differences in genetic variation among geographic regions in *C. briggsae* might allow laboratory approaches, such as quantitative trait loci (QTL) mapping, to identify regions of the genome that are associated with phenotypic differences between strains from different parts of the world. In *C. elegans*, QTLs have been identified for a number of life-history related traits, with evidence of genotype–environment interactions, including in response to temperature (Shook and Johnson, 1999; Knight *et al*., 2001; Ayyadevara *et al*., 2003; Gutteling *et al*., 2007). However, the lack of correlation between biogeographic structure and diversity in *C. elegans* means it is difficult to make inferences about environmental adaptations, and suggests that any QTL identified might not necessarily correlate neatly with ecology.

The biogeographic patterning of *C. briggsae* could make this species more amenable than *C. elegans* to studies of ecological and behavioral adaptation (Baird and Chamberlin, 2006; Cutter *et al*., 2006b). The few studies that examined strain-specific differences in *C. briggsae* have shown that strains from different clades exhibit variation in hybrid compatibilities (Baird, 2002), the patterning of sensory rays in the male tail (Baird, 2001) and vulval cell lineages (Delattre and Félix, 2001). In this study, we demonstrated that additional phenotypic variation is also found within clades, as South African *C. briggsae* strains exhibited a unique pattern of P4.p division despite quite similar molecular sequences to strains from the Tropic of Cancer. It remains to be seen, however, how distinct the genomes of this particular set of strains are, and how much within- and between-clade diversity exists in the species. More samples will also be needed to determine if the latitudinal clade pattern found here is a common global feature, and to more accurately determine how the patterns of migration, outcrossing and selection differ between *C. elegans* and *C. briggsae*. With a growing number of strains being collected from around the world, and increasing genetic and molecular resources, *C. briggsae* presents itself as a useful companion species to *C. elegans*, and might be well suited to studying some questions of gene flow, speciation and adaptation.

Acknowledgements

We are grateful to M Owen-Smith, A Müller and K Amalemba for assistance with field collections, and to C Griffiths and I Gordon for providing laboratory facilities. Worm strains were generously provided by L Carta, K Kiontke and the Caenorhabditis Genetics Center. We thank B Charlesworth and three anonymous reviewers for helpful comments on the manuscript, and to M Ailion, S Baird and D Denver for discussing unpublished results. This work was funded by a University of Edinburgh Development Trust Small Project Grant, and by postgraduate scholarships from the Natural Sciences and Engineering Research Council

(Canada) and the University of Edinburgh School of Biological Sciences to ESD, and by the National Science Foundation (USA) with International Research Fellowship Program grant 0401897 to ADC.

References

Armbruster P, Reed DH (2005). Inbreeding depression in benign and stressful environments. *Heredity* **95**: 235–242.

Ayyadevara S, Ayyadevara R, Vertino A, Galecki A, Thaden JJ, Shmookler Reis RJ (2003). Genetic loci modulating fitness and life span in *Caenorhabditis elegans*: categorial trait interval mapping in CL2a x Beregerac-BO recombinant-inbred worms. *Genetics* **163**: 557–570.

Baïlle D, Barrière A, Félix M-A (2008). *Oscheius tipulae*, a widespread hermaphroditic soil nematode, displays a higher genetic diversity and geographic structure than *Caenorhabditis elegans*. *Mol Ecol*, in press.

Baird SE (2001). Strain-specific variation in the pattern of caudal papillae in *Caenorhabditis briggsae* (Nematoda: Rhabditidae); implications for species identification. *Nematology* **3**: 373–376.

Baird SE (2002). Haldane's rule by sexual transformation in *Caenorhabditis*. *Genetics* **161**: 1349–1353.

Baird SE, Chamberlin HM (2006). *Caenorhabditis briggsae* methods (December 18, 2006). In: The *C. elegans* Research Community (ed). *Wormbook*, doi/10.1895/wormbook.1.128.1, http://www.wormbook.org.

Baird SE, Davidson CR, Bohrer JC (2005). The genetics of ray pattern variation in *Caenorhabditis briggsae*. *BMC Evol Biol* **5**: 3.

Barrière A, Félix M-A (2005). High local genetic diversity and low outcrossing rate in *Caenorhabditis elegans* natural populations. *Curr Biol* **15**: 1176–1184.

Barrière A, Félix M-A (2006). Isolation of *C. elegans* and related nematodes (July 17, 2006). In: The *C. elegans* Research Community (ed). *Wormbook*, doi/10.1895/wormbook.1.115.1, http://www.wormbook.org.

Barrière A, Félix M-A (2007). Temporal dynamics and linkage disequilibrium in natural *Caenorhabditis elegans* populations. *Genetics* **176**: 999–1011.

Charlesworth D (2003). Effects of inbreeding on the genetic diversity of populations. *Philos Trans R Soc Lond B Biol Sci* **358**: 1051–1070.

Coghlan A, Wolfe KH (2002). Fourfold faster rate of genome rearrangement in nematodes than in *Drosophila*. *Genome Res* **12**: 857–867.

Cutter AD (2006). Nucleotide polymorphism and linkage disequilibrium in wild populations of the partial selfer *Caenorhabditis elegans*. *Genetics* **172**: 171–184.

Cutter AD, Baird SE, Charlesworth D (2006a). High nucleotide polymorphism and rapid decay of linkage disequilibrium in wild populations of *Caenorhabditis remanei*. *Genetics* **174**: 901–913.

Cutter AD, Félix M-A, Barrière A, Charlesworth D (2006b). Patterns of nucleotide polymorphism distinguish temperate and tropical wild isolates of *Caenorhabditis briggsae*. *Genetics* **173**: 2021–2031.

de Bono M, Bargmann CI (1998). Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans*. *Cell* **94**: 679–689.

Delattre M, Félix M-A (2001). Polymorphism and evolution of vulval precursor cell lineages within two nematode genera, *Caenorhabditis* and *Oscheius*. *Curr Biol* **11**: 631–643.

Denver DR, Morris K, Thomas WK (2003). Phylogenetics in *Caenorhabditis elegans*: an analysis of divergence and outcrossing. *Mol Biol Evol* **20**: 393–400.

Dolgin ES, Charlesworth B, Baird SE, Cutter AD (2007). Inbreeding and outbreeding depression in *Caenorhabditis* nematodes. *Evolution* **61**: 1339–1352.

Floyd R, Abebe E, Papert A, Blaxter M (2002). Molecular barcodes for soil nematode identification. *Mol Ecol* **11**: 839–850.

Fodor A, Riddle DL, Nelson FK, Golden JW (1983). Comparison of a new wild-type *Caenorhabditis briggsae* with laboratory strains of *Caenorhabditis briggsae* and *C. elegans*. *Nematologica* **29**: 203–217.

Graustein A, Gaspar JM, Walters JR, Palopoli MF (2002). Levels of DNA polymorphism vary with mating system in the nematode genus *Caenorhabditis*. *Genetics* **161**: 99–107.

Gutteling EW, Riksen JAG, Bakker J, Kammenga JE (2007). Mapping phenotypic plasticity and genotype-environment interactions affecting life-history traits in *Caenorhabditis elegans*. *Heredity* **98**: 28–37.

Haag ES, Ackerman AD (2005). Intraspecific variation in *fem-3* and *tra-2*, two rapidly coevolving nematode sex-determining genes. *Gene* **349**: 35–42.

Haber M, Schüngel M, Putz A, Müller S, Hasert B, Schulenburg H (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.

Hillier LW, Miller RD, Baird SE, Chinwalla A, Fulton LA, Koboldt DC *et al*. (2007). Comparison of *C. elegans* and *C. briggsae* genome sequences reveals extensive conservation of chromosome organization and synteny. *PLoS Biol* **5**: e167.

Hodgkin J, Doniach T (1997). Natural variation and copulatory plug formation in *Caenorhabditis elegans*. *Genetics* **146**: 149–164.

Hudson RR (2000). A new statistic for detecting differentiation. *Genetics* **155**: 2011–2014.

Hudson RR, Boos DD, Kaplan NL (1992). A statistical test for detecting geographic subdivision. *Mol Biol Evol* **9**: 138–151.

Hudson RR, Kaplan NL (1985). Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* **111**: 147–164.

Huson DH, Bryant D (2006). Applications of phylogenetic networks in evolutionary studies. *Mol Biol Evol* **23**: 254–267.

Inoue T, Ailion M, Poon S, Kim HK, Thomas JH, Sternberg PW (2007). Genetic analysis of dauer formation in *Caenorhabditis briggsae*. *Genetics* **177**: 809–818.

Jovelin R, Ajie BC, Phillips PC (2003). Molecular evolution and quantitative variation for chemosensory behaviour in the nematode genus *Caenorhabditis*. *Mol Ecol* **12**: 1325–1337.

Kiontke K, Fitch DH (2005). The phylogenetic relationships of *Caenorhabditis* and other rhabditids (August 11, 2005). In: The *C. elegans* Research Community (ed). *Wormbook*, doi/10.1895/wormbook.1.11.1, http://www.wormbook.org.

Kiontke K, Gavin NP, Raynes Y, Roehrig C, Piano F, Fitch DH (2004). *Caenorhabditis* phylogeny predicts convergence of hermaphroditism and extensive intron loss. *Proc Natl Acad Sci USA* **101**: 9003–9008.

Kiontke K, Sudhaus W (2006). Ecology of *Caenorhabditis* species (January 9, 2006). In: The *C. elegans* Research Community (ed). *Wormbook*, doi/10.1895/wormbook.1.37.1, http://www.wormbook.org.

Knight CG, Azevedo RBR, Leroi AM (2001). Testing life-history pleiotropy in *Caenorhabditis elegans*. *Evolution* **55**: 1795–1804.

Koch R, van Luenen HGAM, van der Horst M, Thijssen KL, Plasterk RHA (2000). Single nucleotide polymorphisms in wild isolates of *Caenorhabditis elegans*. *Genome Res* **10**: 1690–1696.

Maupas EF (1900). Modes et formes de reproduction des nématodes. *Arch Zool Expér Gén* **8**: 463–624.

Maydan JS, Flibotte S, Edgley ML, Lau J, Selzer RR, Richmond TA *et al*. (2007). Efficient high-resolution deletion discovery in *Caenorhabditis elegans* by array comparative genomic hybridization. *Genome Res* **17**: 337–347.

Nigon V, Dougherty EC (1949). Reproductive patterns and attempts at reciprocal crossing of *Rhabditis elegans*, and *Rhabditis briggsae* Dougherty and Nigon, 1949 (Nematoda: Rhabditidae). *J Exp Zool* **112**: 485–503.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496–2497.

Schulenburg H, Müller S (2004). Natural variation in the response of *Caenorhabditis elegans* towards *Bacillus thuringiensis*. *Parasitology* **128**: 433–443.

Shook DR, Johnson TE (1999). Quantitative trait loci affecting survival and fertility-related traits in *Caenorhabditis elegans* show genotype-environment interactions, pleiotropy and epistasis. *Genetics* **153**: 1233–1243.

Sokal RR, Rohlf FJ (1995). *Biometry*. W.H. Freeman and Company: New York.

Stein LD, Bao Z, Blasiar D, Blumenthal T, Brent MR, Chen N *et al*. (2003). The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. *PLoS Biol* **1**: e45.

Stewart MK, Clark NL, Merrihew G, Galloway EM, Thomas JH (2005). High genetic diversity in the chemoreceptor superfamily of *Caenorhabditis elegans*. *Genetics* **169**: 1985–1996.

Tajima F (1989). Statistical method for testing the neutral mutation hypothesis. *Genetics* **123**: 585–595.

Viney ME, Gardner MP, Jackson JA (2003). Variation in *Caenorhabditis elegans* dauer larva formation. *Dev Growth Differ* **45**: 389–396.

Supplementary Information accompanies the paper on Heredity website (http://www.nature.com/hdy)