Inversion polymorphisms and nucleotide variability in Drosophila

PETER ANDOLFATTO[1]*, FRANTZ DEPAULIS[1] AND ARCADIO NAVARRO[1]

[1] *Institute of Cell, Animal & Population Biology, University of Edinburgh, Edinburgh, EH9 3JT, UK*

Introduction

Population genetic models of molecular variation usually assume that most genetic variability is neutral or nearly so (Kimura, 1983). However, even if most variability is neutral, patterns of polymorphism may be indirectly shaped by natural selection (Gillespie, 1997). Indeed, theoretical studies show that neutral variability linked to selected sites may be influenced by different types of natural selection, including balancing selection (Strobeck, 1983; Hudson & Kaplan, 1988), Hill-Robertson interference (Hill & Robertson, 1966) or directional selection (Maynard Smith & Haigh, 1974; Charlesworth *et al.*, 1993; Gillespie, 1997). In all these models of evolution, the recombination rate is a key parameter. Other factors being equal, the lower the recombination rate, the larger the chromosomal segment affected by selection at a linked site.

The theoretical importance of recombination is confirmed by what is perhaps one of the most convincing patterns to emerge from population genetics over the last decade: levels of nucleotide polymorphism are reduced in genomic regions with low rates of meiotic crossing-over. Evidence for this trend has been gathered in a wide variety of organisms (Nachman, 1997; Dvorak *et al.*, 1998; Stephan & Langley, 1998; Przeworski *et al.*, 2000). To date, the main body of data comes from *Drosophila melanogaster*, where the local recombination rate accounts for a large portion of the variance in nucleotide diversity among genes (Aquadro *et al.*, 1994). This general pattern is thought to reflect the distinct effects of variation-reducing selection on neutral variability in different recombination environments.

Recombination rates vary both across the genome and between closely related species of *Drosophila* (cf.

Ashburner, 1989; True *et al.*, 1996). In addition, natural populations are known to harbour genetic variation for recombination rate (reviewed by Brooks, 1988). An important class of recombination rate modifiers is represented by chromosomal inversions (Roberts, 1976). First described in *Drosophila* (Sturtevant, 1917), inversions are portions of the chromosomes whose gene order has been reversed relative to the standard reference orientation. Crossing over is strongly suppressed between two paired chromosomes that differ by an inversion. In a population genetic context, inversions would not be important if they were rare in natural populations. In fact, an estimated three quarters of all species in the genus *Drosophila* are polymorphic for inversions (Krimbas & Powell, 1992; Powell, 1997). A large number of these inversions have frequencies in natural populations that exceed several percent. Despite their potential impact on patterns of variability in *Drosophila*, inversions have been largely overlooked in nucleotide variation surveys.

Heterozygosity for an inversion induces a loop structure in paired homologous chromosomes and thus inversions are easy to detect cytologically. This convenient feature motivated the original studies in *Drosophila*. As inversions have no known function, they were first considered to be subject primarily to genetic drift. However, selective effects associated with inversions were soon discovered. Perhaps the most convincing are the patterns of spatial heterogeneity in inversion frequency in natural populations of *Drosophila*. Linkage disequilibrium between some inversions and genetically distant allozyme loci (or characters associated with fitness) suggested that these associations may be selectively maintained. Finally, population cage experiments with inversions by Dobzhansky and colleagues pointed to a selective advantage to heterokaryotypes. Many models have been proposed to account for the maintenance of inversions in natural populations. They generally posit that the inhibition of recombination is advantageous in the presence of some modes of selection,

* Corresponding author: Institute of Cell, Animal & Population Biology, Kings Buildings – West Mains Road, University of Edinburgh, Edinburgh EH9 3JT, Scotland, U.K. Tel: +44 (0) 131 650 8624. Fax: +44 (0) 131 650 6564. email: Peter.Andolfatto@ed.ac.uk.

e.g. epistatic interactions among loci. Readers interested in the above issues should refer to the extensive reviews by Dobzhansky (1970), Krimbas & Powell (1992) and Powell (1997).

This review focuses on the theoretical and empirical research conducted over the past decade, with an emphasis on nucleotide variation patterns in *Drosophila*. Our interest is in the effect of inversions on the level and distribution of nucleotide diversity in natural populations, particularly in relation to simple paracentric inversions (i.e. those not involving the centromere). Our intention is not to review all of the relevant results in detail, but rather to provide examples to illustrate a few, tentative generalisations.

Predicted effects of inversions on recombination and diversity patterns

Inversions have been shown to have several effects on recombination rates (reviewed by Roberts, 1976). First, inversions relocate genes along a chromosome, potentially modifying their recombinational context. An example is a fixed inversion difference between *D. melanogaster* and *D. simulans* on chromosome 3R that relocates a number of genes from a region of normal recombination in *D. simulans* (and its ancestor) to a lower recombination region in *D. melanogaster* (cf. Ashburner, 1989). Second, recombination is reduced by the inhibition of chiasmata by asynapsis in heterokaryotypes (Roberts, 1976; Coyne *et al.*, 1993; Navarro & Ruiz, 1997). This effect is likely to be most pronounced near the breakpoints of paracentric inversions (Novitski & Braver, 1954; Grell, 1962). Finally, recombination is reduced in heterokaryotypes because crossing over events within the inverted region give rise to non-viable aneuploid meiotic products. For paracentric inversions (the most common), most of these products of ovogenesis are eliminated into polar bodies, thus avoiding strong deleterious effects (*Drosophila* males generally lack crossing over).

The main consequence of reduced recombination along the inverted chromosome segment is to subdivide the population into two classes, standard and inverted. These two classes are not completely isolated. Genetic exchanges between chromosomal arrangements within the inverted region (hereafter, gene flux) are still possible, because viable recombinant gametes arise by multiple crossing over (reviewed in Ashburner, 1989) and by gene conversion (Chovnick, 1973). Navarro *et al.* (1997) and Navarro & Ruiz (1997) studied the effect of single inversion differences on gene flux between paired homologous chromosomes. The authors provide equations that predict the recombination rate at every site along the inverted chromosome, given the physical and genetic length of

an inversion. Three conclusions emerge. First, the reduction in recombination rate is expected to be stronger for shorter inversions. Second, the major effect of inversions within the inverted segment is to reduce the rate of crossing over but not that of gene conversion. The suppression of crossing over may be large enough for gene conversion to dominate as the main mode of gene flux, even over large physical distances. The third, and most important, conclusion is that the reduction in recombination is not expected to be uniform along the inverted chromosome. Thus, inversions not only reduce recombination in inversion heterozygotes, but also redistribute recombination events. They reduce gene flux between arrangements to almost zero near the breakpoints, but rates of exchange remain high in central regions.

Navarro *et al.* (2000) modelled inversions as balanced polymorphisms. The authors conclude that the effect of inversions on nucleotide variability of linked regions depends mainly on two factors: gene flux and the age of the inversion polymorphism. An example to illustrate the predicted effects is presented in Fig. 1. Here, an inversion appears at some time in the past and rises immediately to an equilibrium frequency of 50%. The rapid rise in frequency leads to a bottleneck among standard chromosomes, and hence to a gradual loss of variability within the standard class. However, gene flux between arrangements causes transient losses and gains of variability within both classes. The interaction of these two processes can produce complex patterns of diversity.

In general, if the rate of gene flux is on the order of the reciprocal of the effective population size (N_e) or higher (Fig. 1 and results not shown), the effects of an inversion on patterns of diversity in the standard arrangement, and on the population as a whole, are relatively minor. For large paracentric inversions in *Drosophila*, only regions very close to inversion breakpoints (i.e. a small proportion of the genome) will be markedly affected. For example, in Fig. 1, diversity is (transiently) reduced by only 25%. The main effects on nucleotide diversity levels and the frequency spectrum of mutations will be seen in the inverted arrangement. Diversity levels within this class are markedly reduced, as expected after a partial selective sweep. In this model, the inverted class reaches its equilibrium frequency with no variation; it then gains diversity through mutation and gene flux. New mutations will be at low frequency, but these variants will be few relative to the higher frequency mutations imported by gene flux. Thus, despite drastic changes in the frequency of the inverted class, the frequency spectrum of mutations will not be markedly distorted, save for regions with flux levels much lower than $1/N_e$ (Fig. 1 and results not shown). Note that the model assumes a single inversion difference; more complex arrangements may show effects on diversity
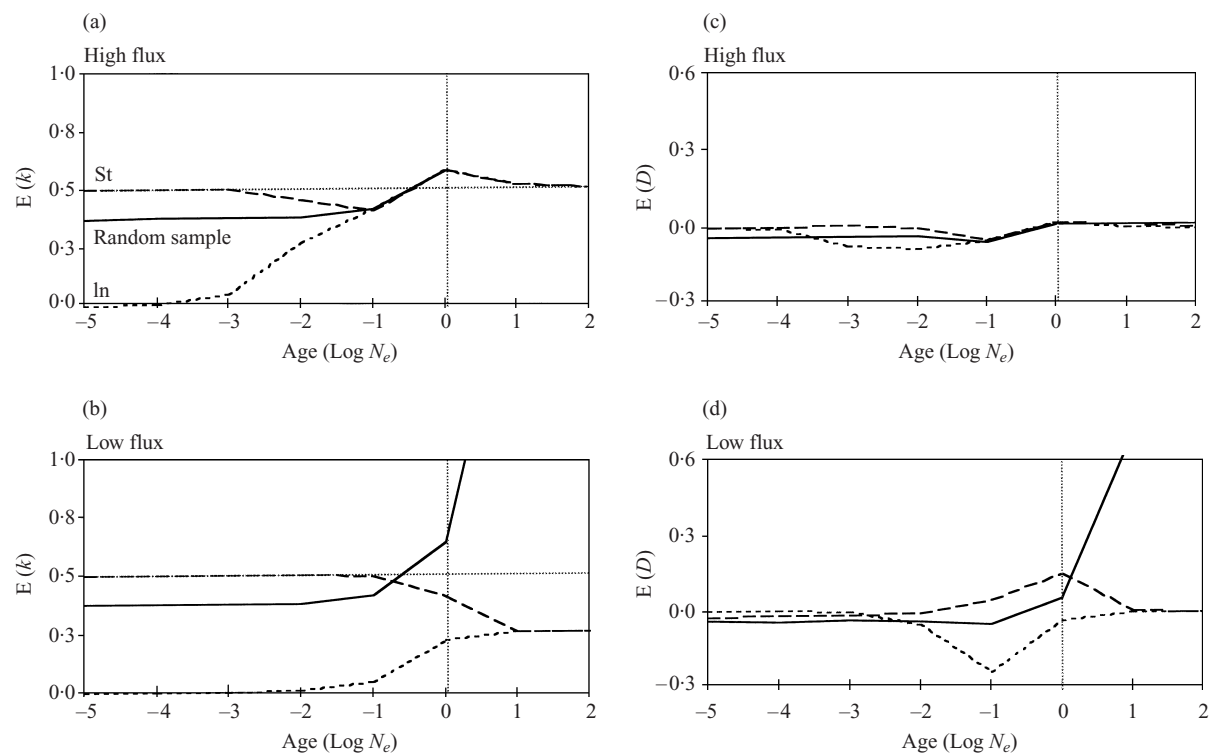
Fig. 1. Changes in expected average pairwise diversity, E(k), and a measure of the skew in the frequency spectrum of mutations, E(D), (Tajima 1989) over time in a sample of 25 alleles taken at random from the whole population (*random sample*, solid line), the standard arrangement (*St*, dashed line) or the inverted arrangement (*In*, dotted line). Inversion and Standard chromosomes are assumed to be maintained at stable 1:1 frequencies by strong balancing selection. Results are based on 100 000 runs of the program described in Navarro *et al.* (2000) with the difference that, here, the inversion is assumed to increase instantaneously in frequency. Time is measured in units of Log(N_e generations) since the origin of the inversion polymorphism, where N_e is the effective population size of the species. Parameter values: $N_e = 10^6$, $\mu = 1.25 \times 10^{-7}$; $\theta = 0.5$; gene flux in heterokaryotypes is high in (a) and (c) (10^{-4} per generation), and low in (b) and (d) (10^{-10} per generation).

and the frequency spectrum of mutations over larger distances.

The short-term effect (i.e. fewer than N_e generations) of a balanced inversion system will be to reduce overall levels of variability. If, in contrast, an inversion polymorphism has been maintained by selection at a constant frequency for a very long time (i.e. $> 10N_e$ generations), the effect on total variability will be quite different (Fig. 1). A long-lived balanced polymorphism is analogous to population subdivision (Strobeck, 1983). The two arrangement classes will diverge over time. Differentiation between arrangements will be highest around the breakpoints and should decrease rapidly with increasing distance from them (Strobeck, 1983; Navarro *et al.*, 2000). Over much of the length of the inversion, total nucleotide variability, and the frequency spectrum in the population as a whole, will be close to that expected in a panmictic neutral population, if gene flux is high enough to homogenise the two arrangements (i.e. on the order of $1/N_e$ or higher). For large paracentric inversions in *Drosophila*, this condition is likely to be met. Thus, only loci very closely linked to inversion breakpoints are expected to have elevated total

variability and a markedly skewed frequency spectrum (see Fig. 1).

These alternative inversion age scenarios illustrate that the overall effect of inversions on nucleotide variability is determined primarily by levels of gene flow and the inversion's average sojourn time (or 'age') in the population. The precise interpretation of 'age' depends on the model. Roughly, if the frequency of inversions in natural populations is determined by selection (as modelled by Navarro *et al.*, 2000), the 'age' of an inversion can be interpreted as the last time at which the inversion experienced a rapid increase in frequency. In reality, inversions are likely to have more complicated historical frequency trajectories. The quantitative effects may be different from those modelled, but the qualitative effects should be much the same.

The simplest assumption, and the point of departure for most population genetic investigations, is that inversions have drifted (neutrally) to their present frequency. However unlikely, a neutral model for inversions makes some predictions that overlap with those of selection models. For example, Innan & Tajima (1997) studied the properties of allelic classes

in neutral genealogies with no recombination. In this model, a single mutation distinguishes two allelic sub-genealogies. A corollary of their work is that the two allelic classes (e.g. defined by the inversion mutation) will show considerable differentiation and reduced variability, even if the inversion is evolving neutrally. The class most strongly affected is predicted to be the derived one (in this case, the inversion). The allelic subsample effects described by Innan & Tajima (1997) are expected to dissipate rapidly with gene flux between arrangements. Thus, a potential way to distinguish between a selective increase and neutral evolution is to investigate how quickly variability at loci within the inverted arrangement recovers with genetic distance from the inversion breakpoint.

Empirical patterns

To date, there are six inversion systems in *Drosophila* for which data are available for linked loci and for which inverted alleles have been determined: the *sex ratio* complex (Babcock & Anderson, 1996; Kovacevic & Schaeffer, 2000) and the third chromosome inversions of *D. pseudoobscura* (Aquadro *et al.*, 1991; Popadic *et al.*, 1995); the O chromosome inversion system of *D. subobscura* (Munté *et al.*, 2000; Rozas & Aguadé, 1994; Rozas *et al.*, 1999), the *2j* inversion complex of *D. buzzatii* (Cáceres *et al.*, 1999) and *In(3L)Payne* (Wesley & Eanes, 1994; Hasson & Eanes, 1996) and *In(2L)t* of *D. melanogaster* (Aguadé, 1988, 1998, 1999; Bénassi *et al.*, 1993, 1999; Veuille *et al.*, 1998; Andolfatto, *et al.*, 1999; Depaulis, *et al.*, 1999, 2000; Andolfatto & Kreitman, 2000). Here, we focus on *D. melanogaster*, where most of these studies have been conducted. This particular model system has several advantages. *D. melanogaster* is polymorphic for at least 339 inversions (Lemeunier & Aulard, 1992). Four of these, one on each major autosomal arm, have cosmopolitan distributions and reach high frequencies in equatorial populations. Perhaps the most useful feature of this species is the availability of detailed physical and genetic maps (cf. Heino, 1994; Flybase: http://flybase.bio.indiana. edu), including several estimates of flux between arrangements (Payne, 1924; Chovnik, 1973; Malpica *et al.*, 1987). This information provides the physical and genetic position of loci in relation to inversion breakpoints. In addition, *In(2L)t* and *In(3L)P* are simple inversions, each having only two breakpoints (unlike the *sex ratio* inversions of *D. pseudoobscura*, the O inversion complex of *D. subobscura*, or the *2j* complex of *D. buzzatii*, reviewed in Krimbas & Powell, 1992). This feature allows for simpler predictions about their likely effects on recombination.

What do the molecular data tell us about levels of gene flux between arrangements? Estimates of gene flux are not independent of assumptions about the underlying population genetic model. For example, a selective sweep (e.g. Depaulis *et al.*, 2000) or a recent rapid change in inversion frequency (Andolfatto *et al.*, 1999) would considerably weaken the power to detect exchange even if the rate is high. Several generalisations can nonetheless be made. First, the molecular analysis of regions spanning inversion breakpoints of three common inversions in *Drosophila* indicate that they are unique-event mutations (Wesley & Eanes, 1994; Andolfatto *et al.*, 1999; Cáceres *et al.*, 1999). Second, breakpoint regions generally reveal fixed differences between arrangements, but few or no shared polymorphisms, while the opposite pattern is observed at loci in the middle of (or far outside) the inverted region. This pattern is consistent with reduced gene flux near breakpoints relative to loci far from them (Hasson & Eanes, 1996; Andolfatto *et al.*, 1999; Kovacevic & Schaeffer, 2000). The one exception is the O complex inversion system, where the position of the marker locus in relation to inversion breakpoints varies according to which karyotypes are considered (Rozas *et al.*, 1999). Thus, there appears to be considerable gene flux between arrangements (over much of their length) and even complex inversion systems, such as *sex ratio* or the O complex (Rozas & Aguadé, 1994; Kovacevic & Schaeffer, 2000), are not genetically congealed blocks. This is not to say that multiple sites within an inversion are not under selection, or interacting epistatically, as envisaged by Dobzhansky (1970). However, in the face of appreciable levels of genetic exchange, only neutral loci tightly linked to the selected loci are expected to reveal the footprint of selection (Ishii & Charlesworth, 1977; Strobeck, 1983; Hudson & Kaplan, 1988).

Excluding small regions immediately spanning inversion breakpoints, the rate of gene flux between arrangements is likely to be several orders of magnitude higher than the rate of mutation (Payne, 1924; Chovnick, 1973; Malpica *et al.*, 1987; Navarro *et al.*, 1997). In the case of the relatively large and simple inversions *In(2L)t* and *In(3L)P*, gene flux is expected to be less than 10^{-6} (per generation) only within a few hundred kilobases of the breakpoints. As a result, the pattern of nucleotide variation at random loci will not be informative about the mode of selection acting on inversions, or about their population history. Studies have therefore focused on loci near inversion breakpoints (Wesley & Eanes, 1994; Babcock & Anderson, 1996; Andolfatto *et al.*, 1999; Cáceres *et al.*, 1999), where the homogenising effects of gene flux between arrangements are likely to be the weakest.

How old are common inversion polymorphisms in *Drosophila*? As outlined above, the effect of a selected inversion on neutral variation depends on the inversion's frequency over time and the timing of its putative rise in frequency. Approaches to estimating the 'age' of inversions have focused on levels of
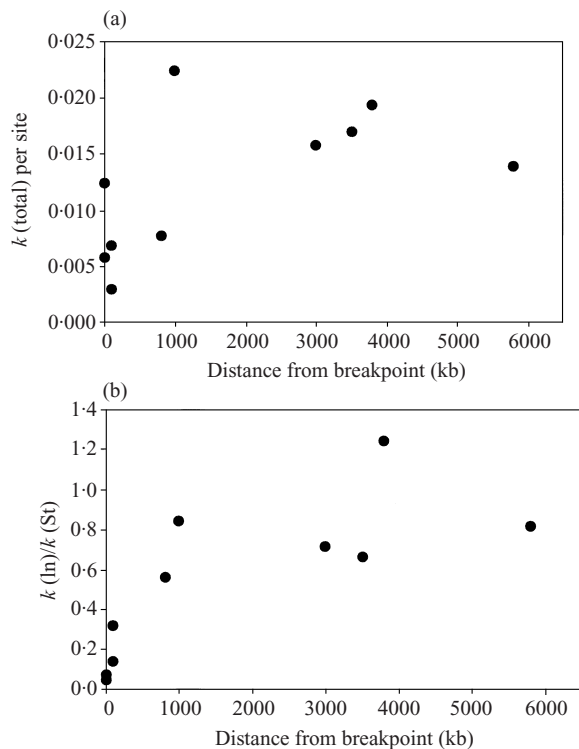
Fig. 2. Change in the average pairwise diversity at silent sites with distance (kilobases) from the inversion breakpoint. Data come from the following sources: *In(2L)t*: Breakpoint (Andolfatto *et al.*, 1999), *Acp26Aab* (Aguadé, 1998), *Acp29AB* (Aguadé, 1999), *Vha* (Depaulis, *et al.*, 2000), *Fbp2* (Bénassi *et al.*, 1998), *Adh* (S.-C. Tsaur, unpublished data) and *Su(H)* (Depaulis *et al.*, 1999); *In(3L)P*: Breakpoint, *Est-6* and *Hsp83* (Hasson & Eanes, 1996). (a) Variability for the whole sample (Spearman's rank correlation $p = 0.03$) (b) Variability for *In* chromosomes. In order to correct for individual locus effects, *In* variability is scaled by *St* variability (Spearman's rank correlation $p < 0.001$).
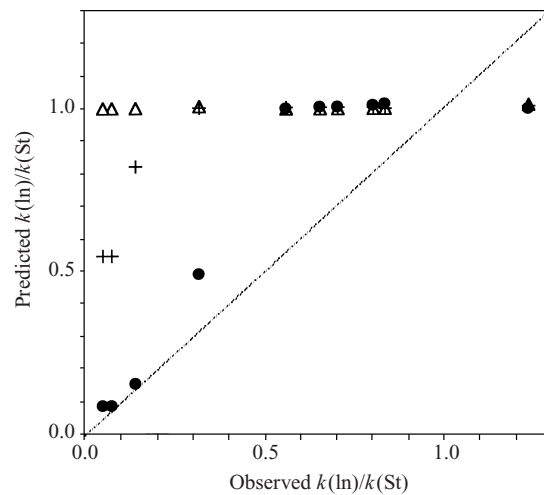

Fig. 3. Observed average pairwise diversity at silent sites (*In/St*, as in Fig. 2b) *vs.* predicted values. Predicted variability is based on an hypothetical system where the two arrangements instantaneously reach 1:1 frequencies and remain stable (Navarro *et al.*, 2000, with the same difference as in Fig. 1). Recombination rates were estimated following the approaches of Navarro *et al.* (1997) and Navarro & Ruiz (1997). Estimates of gene flux in *In(2L)t* and *In(3L)P* heterozygotes are based on measurements by Malpica *et al.* (1987) and Payne (1924), respectively. Circles correspond to the predictions for a young balanced polymorphism (established 10^5, or $0.1N_e$ generations ago), crosses correspond to a middle-aged polymorphism (10^6, or N_e generations), and triangles to an old polymorphism (10^7, or $10N_e$ generations). Qualitatively, these results are robust to a wide range of assumed inversion frequencies.

genetic differentiation between chromosomal types (Aquadro *et al.*, 1991; Rozas & Aguadé, 1994; Babcock & Anderson, 1996; Hasson & Eanes, 1996; Andolfatto *et al.*, 1999; Kovacevic & Schaeffer, 2000). Given the geographic and temporal variation in inversion frequencies in *Drosophila* populations (Krimbas & Powell, 1992), estimates of age based on allele frequencies (e.g. Griffiths & Tavaré, 1998) are expected to be unreliable. However, estimates of the age from levels of differentiation also depend on the model of inversion evolution. An alternative is to assume a specific selection model, such as balancing selection (Andolfatto *et al.*, 1999) or recent directional selection (Depaulis *et al.*, 2000). This said, all the methods yield inversion age estimates that are on the order of N_e generations or fewer. Thus, the available data suggest that the effect of inversions is more often to reduce nucleotide variability in natural populations than to increase it.

Nucleotide variation studies at multiple loci linked to *D. melanogaster* inversions *In(3L)P* and *In(2L)t*

have examined patterns of variability and differentiation between arrangements, as a function of distance from inversion breakpoints (Wesley & Eanes, 1994; Hasson & Eanes, 1996; Aguadé, 1998, 1999; Andolfatto *et al.*, 1999; Bénassi *et al.*, 1999; Depaulis *et al.*, 1999, 2000). In Fig. 2, we plot a measure of nucleotide diversity within inversions for ten loci as a function of physical distance from the inversion breakpoints. As expected under a recent balanced polymorphism model, there is a trend towards higher diversity within inverted chromosomes as one moves away from the breakpoint (Fig. 2b, $p < 0.001$ by a two tailed Spearman rank correlation test). As can be seen, inversion effects on levels of diversity may extend as far as 1000 kilobases. A marginally significant correlation is observed between total diversity and distance (Fig. 2a, $p = 0.03$ by a two-tailed Spearman rank correlation test). Interestingly, under a neutral model of inversion evolution, reduced *total* diversity is not expected at the breakpoint (cf. Innan & Tajima, 1997).

In Fig. 3, we compare the data to predictions of the balanced polymorphism model of inversions. As been seen, a recent balanced polymorphism provides a better fit to the data than an old balanced

polymorphism. A neutral history for the inversion (i.e. Innan & Tajima, 1997) cannot be excluded on the basis of the data. A subset of data points (namely the two breakpoint regions themselves, as well as *Hsp83* and *Vha*) are in the range of flux rates that might allow one to distinguish between a recent selected rise in inversion frequency and a neutral inversion. However, the extent to which the allelic subsample effects (Innan & Tajima, 1997) can account for the observed reductions in diversity depend on the frequency of the inversion and the extent to which gene flux is suppressed, the parameter values of which are uncertain. Thus, the only firm conclusion is that these two inversions are unlikely to be ancient polymorphisms. This is surprising given the independent evidence for selection acting on inversions (reviewed by Krimbas and Powell, 1992). One possibility is that inversion frequencies are determined by episodes of directional selection (including short-lived balancing selection) rather than long-lived balancing selection.

Positive selection and inversion polymorphisms

If positive directional selection is frequent in natural populations, how inversions remain polymorphic is an enigma. Indeed, strong selection should either fix inversions in the population, or cause their loss (depending on the inversion's frequency). Yet, there is only one fixed inversion difference between *D. melanogaster* and *D. simulans* (Ashburner, 1989), while many inversions are segregating at appreciable frequencies in *D. melanogaster*. One possibility is that the fixation of advantageous variants may be infrequent relative to the turnover rate of inversions. A second possibility is that positive selection is usually weaker than either the rate of gene flux between arrangements or the strength of selection maintaining inversions.

Given that inversions are present at intermediate frequencies in some populations, it is unclear what patterns of nucleotide diversity to expect from directional selection; these might be quite complex. An example is the heterogeneity in patterns recovered at multiple loci linked to *In(2L)t* in African populations of *D. melanogaster*. For these populations, *In(2L)t* frequencies are in the neighbourhood of 50% (Aguadé, 1998, 1999; Veuille *et al.*, 1998; Bénassi *et al.*, 1999; Depaulis *et al.*, 1999, 2000; Andolfatto & Kreitman, 2000). Departures from neutral expectations (consistent with recent positive selection) were found at three loci (*Su(H)*, *Fbp2* and *Vha*) but not at several others (*Adh*, *Acp26Aab*, *Acp29AB* and *In(2L)t* proximal breakpoint). At present, the results are difficult to interpret. In particular, the predicted level of gene flux cannot easily be related to the diversity patterns recovered at these five loci (i.e. gene flux is low for *Vha* and the *In(2L)t* breakpoint sequences,

but relatively high for all others). However, in the case of *Su(H)*, *Fbp2* and *Vha* (sampled from the same population), simple demographic explanations can be ruled out (Galtier *et al.*, 2000).

If a population is genetically subdivided by an inversion, directional selection may increase differentiation between standard and inverted arrangements (Slatkin & Wiehe, 1998; Depaulis *et al.*, 1999), even at loci far from the breakpoints. Furthermore, restricted gene flux may slow the decay of linkage disequilibrium generated by a partial selective sweep. Theoretical models of the effects of positive directional selection on various aspects of polymorphism data (i.e. diversity levels, the frequency spectrum of mutations and patterns of linkage disequilibrium) may have to account for the genetic population structure introduced by inversions if they are to be meaningfully related to data from species polymorphic for inversions.

Conclusions

Our comparison of theoretical predictions and empirical patterns suggests that inversions are generally not ancient balanced polymorphisms. On the contrary, a wide variety of approaches with different inversion systems suggest ages that are lower than or comparable to the neutral time scale of N_e generations. Given these relatively young estimates of inversion age, the average lifespan of an inversion will generally not be long enough to increase variability in a population. In fact, it is expected that the overall effect of inversions on total population variability will be a reduction over much of the length of the chromosomal regions polymorphic for inversions. The magnitude of this predicted reduction depends on the level of gene flux. For the common cosmopolitan inversions of *D. melanogaster*, the reduction in diversity, and the effect on the frequency spectrum, are expected to be relatively minor. It should be noted, however, that it is difficult to account for the cumulative effects of the hundreds of low frequency inversions segregating in *D. melanogaster* populations. In the case of the more complex inversion systems discussed, quantitative predictions are even more difficult.

The rate of recombination is an important determinant of expected patterns of genome-wide nucleotide variation in *Drosophila*. Common polymorphic inversions, through their effects on recombination, make the interpretation of patterns of variation at many loci in *Drosophila* uncertain. Not only are inversions expected to have several direct effects on levels of diversity, but their presence will also modify the effects of directional selection on linked loci. Short of choosing a system with few or no common inversions, a comparative approach may be informative. For example, to assess the impact of inversions

on diversity, one might compare patterns of variation in *D. melanogaster* to those of its close relative *D. simulans*, which has few inversions and thus constitutes a 'control' (e.g. Andolfatto, 2001). Comparing populations with different inversion frequencies may also be informative (e.g. Veuille *et al.*, 1998; Andolfatto & Kreitman, 2000), although the histories of two populations of the same species cannot be considered to be independent. To avoid differences in life histories or demographic histories, it may be preferable to compare chromosomal arms within a species or population. Potentially useful model species, in this respect, are *D. pseudoobscura* and members of the *D. repleta* group (reviewed in Krimbas & Powell, 1992), both of which have some chromosomal arms that are relatively free of inversions.

References

Aguadé, M. (1988). Restriction map variation at the *Adh* locus of *Drosophila melanogaster* in inverted and non-inverted chromosomes. *Genetics* **119**, 135–140.

Aguadé, M. (1998). Different forces drive the evolution of the *Acp26Aa* and *Acp26Ab* accessory gland genes in the *Drosophila melanogaster* species complex. *Genetics* **150**, 1079–1089.

Aguadé, M. (1999). Positive selection drives the evolution of the *Acp29AB* accessory gland protein in *Drosophila*. *Genetics* **152**, 543–551.

Andolfatto, P. (2001). Contrasting patterns of X-linked and autosomal nucleotide variation in *Drosophila melanogaster* and *D. simulans*. *Molecular Biology and Evolution* **18**, 291–298.

Andolfatto, P. & Kreitman, M. (2000). Molecular variation at the *In(2L)t* proximal breakpoint site in natural populations of *Drosophila melanogaster* and *D. simulans*. *Genetics* **154**, 1681–1691.

Andolfatto, P., Wall, J. D. & Kreitman, M. (1999). Unusual haplotype structure at the proximal breakpoint of *In(2L)t* in a natural population of *Drosophila melanogaster*. *Genetics* **153**, 1297–1311.

Aquadro, C. F., Begun, D. J. & Kindahl, E. C. (1994). Selection, recombination and DNA polymorphism in *Drosophila*. In *Non-neutral Evolution* (ed. B. Golding), pp. 46–56. New York: Chapman & Hall.

Aquadro, C. F., Weaver, A. L., Schaeffer, S. W. & Anderson, W. W. (1991). Molecular evolution of inversions in *Drosophila pseudoobscura*: the amylase gene region. *Proceedings of the National Academy of Sciences of the USA* **88**, 305–309.

Ashburner, M. (1989). *Drosophila. A Laboratory Handbook*. New York: Cold Spring Harbor.

Babcock, C. S. & Anderson, W. W. (1996). Molecular evolution of the Sex-Ratio inversion complex in *Drosophila pseudoobscura*: analysis of the *Esterase-5* gene region. *Molecular Biology and Evolution* **13**, 297–308.

Bénassi, V., Aulard, S., Mazeau, S. & Veuille, M. (1993). Molecular variation of *Adh* and *P6* genes in an African population of *Drosophila melanogaster* and its relation to chromosomal inversions. *Genetics* **134**, 789–799.

Bénassi, V., Depaulis, F., Meghlaoui, G. K. & Veuille, M. (1999). Partial sweeping of variation at the *Fbp2* locus in a west African population of *Drosophila melanogaster*. *Molecular Biology and Evolution* **16**, 347–353.

Brooks, L. D. (1988). The evolution of recombination rates. In *The Evolution of Sex*: *an Examination of Current Ideas* (ed. R. E. Michod & B. R. Levin), pp. 87–105. Sunderland, MA: Sinauer and Associates.

Cáceres, M., Ranz, J. M., Barbadilla, A., Long, M. & Ruiz, A. (1999). Generation of a widespread *Drosophila* inversion by a transposable element. *Science* **285**, 415–418.

Charlesworth, B., Morgan, M. T. & Charlesworth, D. (1993). The effect of deleterious alleles on neutral molecular variation. *Genetics* **134**, 1289–1303.

Chovnick, A. (1973). Gene conversion and transfer of genetic information within the inverted region of inversion heterozygotes. *Genetics* **73**, 123–131.

Coyne, J. A., Meyers, W., Crittenden, A. P. & Sniegowski, P. (1993). The fertility effects of pericentric inversions in *Drosophila melanogaster*. *Genetics* **134**, 487–496.

Depaulis, F., Brazier, L., Mousset, S., Turbé, A. & Veuille, M. (2000). Selective sweep near the *In(2L)t* breakpoint in an African population of *Drosophila melanogaster*. *Genetical Research* **76**, 149–158.

Depaulis, F., Brazier, L. & Veuille, M. (1999). Selective sweep at the *Drosophila melanogaster Suppressor of Hairless* locus and its association with the *In(2L)t* inversion polymorphism. *Genetics* **152**, 1017–1024.

Dobzhansky, T. (1970). *Genetics of the Evolutionary Process*. New York: Columbia University Press.

Dvorak, J., Luo, M. C. & Yang, Z. L. (1998). Restriction fragment length polymorphism and divergence in the genomic regions of high and low recombination in self-fertilizing and cross-fertilizing aegilops species. *Genetics* **148**, 423–434.

Galtier, N., Depaulis, F. & Barton, N. H. (2000). Detecting bottlenecks and selective sweeps from DNA sequence polymorphism. *Genetics* **155**, 981–987.

Gillespie, J. H. (1997). Junk ain't what junk does: neutral alleles in a selected context. *Gene* **205**, 291–299.

Grell, R. (1962). A new model for secondary nondisjunction: the role of distributive pairing. *Genetics* **47**, 1737–1754.

Griffiths, R. C. & Tavaré, S. (1998). The age of a mutant in a general coalescent tree. *Stochastic Models* **14**, 273–295.

Hasson, E. & Eanes, W. F. (1996). Contrasting histories of three gene regions associated with *In(3L)Payne* of *Drosophila melanogaster*. *Genetics* **144**, 1565–1575.

Heino, T. I., Saura, A. O. & Sorsa, V. (1994). Maps of the salivary gland chromosomes of *Drosophila melanogaster*. *Drosophila Information Service* **73**, 621–738.

Hill, W. G. & Robertson, A. (1966). The effect of linkage on the limits of artificial selection. *Genetical Research* **8**, 269–294.

Hudson, R. R. & Kaplan, N. L. (1988). The coalescent process in models with selection and recombination. *Genetics* **120**, 831–840.

Innan, H. & Tajima, F. (1997). The amounts of nucleotide variation within and between allelic classes and the reconstruction of the common ancestral sequence in a population. *Genetics* **147**, 1431–1444.

Ishii, K. & Charlesworth, B. (1977). Associations between allozyme loci and gene arrangements due to hitch-hiking effects of new inversions. *Genetical Research* **30**, 93–106.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge, UK: Cambridge University Press.

Kovacevic, M. & Schaeffer, S. W. (2000). Molecular population genetics of X-linked genes in *Drosophila pseudoobscura*. *Genetics* **156**, 155–172.

Krimbas, C. B. & Powell, J. R. (1992). *Drosophila Inversions Polymorphism*. Cleveland: CRC Press.

Lemeunier, F. & Aulard, S. (1992). Inversion polymorphism in *Drosophila melanogaster*. In *Drosophila Inversion Polymorphism* (ed. C. B. Krimbas & J. R. Powell), pp. 339–405. Boca Raton, FL: CRC Press.

Malpica, J. M., Vassalo, J. M., Frias, A. & Fuentes-Bol, F. (1987). On recombination among *In(2L)t*, *alpha Gpdh* and *Adh* in *Drosophila melanogaster*. *Genetics* **133**, 141–142.

Maynard Smith, J. & Haigh, J. (1974). The hitch-hiking effect of favorable gene. *Genetical Research* **23**, 23–35.

Munté, A., Aguadé, M. & Segarra, C. (2000). Nucleotide variation at the *yellow* gene region is not reduced in *Drosophila subobscura*: a study in relation to chromosomal polymorphism. *Molecular Biology and Evolution* **17**, 1942–1955.

Nachman, M. W. (1997). Patterns of DNA variability at X-linked loci in *Mus domesticus*. *Genetics* **147**, 1303–1316.

Navarro, A., Barbadilla, A. & Ruiz, A. (2000). Effect of inversion polymorphism on the neutral nucleotide variability of linked chromosomal regions in *Drosophila*. *Genetics* **155**, 685–98.

Navarro, A., Betran, E., Barbadilla, A. & Ruiz, A. (1997). Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. *Genetics* **146**, 695–709.

Navarro, A. & Ruiz, A. (1997). On the fertility effects of pericentric inversions. *Genetics* **147**, 931–933.

Novitski, E. & Braver, G. (1954). An analysis of crossing over within an heterozygous inversion in *Drosophila melanogaster*. *Genetics* **39**, 197–209.

Payne, F. (1924). Crossover modifiers on the third chromosome of *Drosophila melanogaster*. *Genetics* **9**, 327.

Popadic, A., Popadic, D. & Anderson, W. W. (1995). Interchromosomal exchange of genetic information between gene arrangements on the third chromosome of *Drosophila* pseudoobscura. *Molecular Biology and Evolution* **12**, 938–943.

Powell, J. (1997). *Progress and Prospects in Evolutionary Biology: the Drosophila Model*. New York: Oxford University Press.

Przeworski, M., Hudson, R. R. & Di Rienzo, A. (2000). Adjusting the focus on human variation. *Trends in Genetics* **16**, 296–302.

Roberts, P. A. (1976). The genetics of chromosome abberation. In *The Genetic and Biology of Drosophila* (ed. M. Ashburner & E. Novitski), pp. 67–184. London and New York: Academic Press.

Rozas, J. & Aguadé, M. (1994). Gene conversion is involved in the transfer of genetic information between naturally occurring inversions of *Drosophila*. *Proceedings of the National Academy of Sciences of the USA* **91**, 11517–11521.

Rozas, J., Segarra, C., Ribo, G. & Aguadé, M. (1999). Molecular population genetics of the *rp49* gene region in different chromosomal inversions of *Drosophila subobscura*. *Genetics* **151**, 189–202.

Slatkin, M. & Wiehe, T. (1998). Genetic hitchhiking in a subdivised population. *Genetical Research* **71**, 155–160.

Stephan, W. & Langley, C. H. (1998). DNA polymorphism in *Lycopersicon* and crossing-over per physical length. *Genetics* **150**, 1585–1593.

Strobeck, C. (1983). Expected linkage disequilibrium for a neutral locus linked to a chromosomal arrangement. *Genetics* **103**, 545–555.

Sturtevant, A. H. (1917). Genetic factors affecting the strength of linkage in *Drosophila*. *Proceedings of the National Academy of Sciences of the USA* **3**, 555.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595.

True, J. R., Mercer, J. M. & Laurie, C. C. (1996). Differences in crossover frequency and distribution among three sibling species of *Drosophila*. *Genetics* **142**, 507–523.

Veuille, M., Bénassi, V., Aulard, S. & Depaulis, F. (1998). Allele-specific population structure of *Drosophila melanogaster alcohol dehydrogenase* at the molecular level. *Genetics* **149**, 971–981.

Wesley, C. S. & Eanes, W. F. (1994). Isolation and analysis of the breakpoint sequences of chromosome inversion *In(3L)Payne* in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the USA* **91**, 3132–3126.