Genet. Res., Camb. (2001), **77**, *pp.* 261–275. *With 5 figures.* Printed in the United Kingdom

Transposable element number in mixed mating populations

MARTIN T. MORGAN*
School of Biological Sciences, Washington State University, Pullman, WA 99164–4236, USA

(*Received 22 May 2000 and in revised form 9 January 2001*)

Summary

Theoretical population genetic studies of transposable elements focus almost exclusively on random mating species, whereas many plants reproduce through partial or substantial self-fertilization. Here I develop computer simulation and analytic approximations of simplified element dynamics (transposition balanced by selective elimination) in partially self-fertilizing populations, using Ty1-*copia* elements for biological inspiration. Under the most plausible models and parameter values, element numbers decrease with self-fertilization when element insertions are deleterious, but may increase when ectopic exchange regulates element number. Conclusions for models of ectopic exchange depend in part on parameters for which little firm empirical evidence is available. Small changes in selfing rate can lead to abrupt changes in element number when homozygous and heterozygous elements have markedly different fitness effects. Equilibrium element numbers can be sensitive to population size, especially at high selfing rates. Elements are frequently lost in small highly selfing populations under the deleterious insertion model. In contrast, small highly selfing populations can accumulate very large numbers of elements under ectopic exchange. Empirical data on element number and localization in plants with different mating systems suggests that deleterious insertion, rather than ectopic exchange, may regulate element number. Limitations to available empirical data, especially the lack of comparison between closely related species differing in mating system, mean that this conclusion is tentative.

1. Introduction

Mobile genetic elements (e.g. retroposons) are ubiquitous components of plant genomes, as evidenced by PCR methods that detect retroposons in many, diverse plant species (Flavell *et al.*, 1992*a*; Voytas *et al.*, 1992; Hirochika & Hirochika, 1993; Suoniemi *et al.*, 1998; Kumar & Bennetzen, 1999). Retroposons account for considerable amounts of genomic DNA (e.g. $> 1/2$ the nuclear DNA of maize; SanMiguel *et al.*, 1996), with copy numbers ranging from tens to millions per plant (Flavell *et al.*, 1992*a*). Evidence indicates that retroposons participate in phenotypic mutation (Berg & Howe, 1989; Bhattacharyya *et al.*, 1990), quantitative and fitness variation (Charlesworth, 1987; Mackay *et al.*, 1992; Keightley, 1994; Lyman *et al.*, 1996), gene regulation (Weil & Wessler, 1990; White *et al.*, 1994; Wessler *et al.*, 1995), deleterious and beneficial mutation (McDonald, 1995), adaptive host evolution (McClintock, 1984; Wilke *et al.*, 1992; White *et al.*, 1994), response to stress (Wessler, 1996; Mhiri *et al.*, 1997; Capy *et al.*, 2000) and genomic structure (Nevers & Saedler, 1977; Bennetzen, 1996; Voytas, 1996; Matzke & Matzke, 1998). Element copy number varies greatly between plant species (Grandbastien, 1992): *Arabidopsis* has until recently been thought to contain only a few apparently non-functional elements (Voytas *et al.*, 1990), whereas maize and barley may have hundreds of thousands of copies (Pearce *et al.*, 1996*b*). Such variation in copy number contrasts with the situation in insects and mammals where, with the exception of mammalian LINEs, elements are typically represented by tens of copies (Kumar, 1996). Despite the wealth of empirical data and diversity of selective explanations (Kidwell & Lisch, 2000), few studies develop theoretical expectations for element numbers in species, such as

* Tel: +1 (509) 335 5641. Fax: +1 (509) 335-3517.
e-mail: mmorgan@wsu.edu

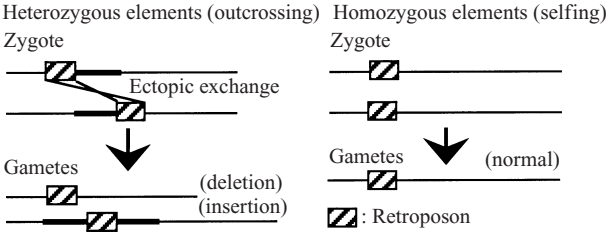

Fig. 1. Ectopic exchange reduces fitness through production of insertions, deletions and translocations. Heterozygosity may increase ectopic exchange (left) compared with homozygosity (right).

many plants, with reproduction through means other than random mating.

The current population genetic paradigm for understanding element dynamics involves a balance between transposition and selection (Charlesworth & Langley, 1989; Charlesworth *et al*., 1994). Each element replicates with probability u, and is excised with probability v. In the ith individual, with y_i elements, the rate of transposition is $(u-v)y_i$. Transposition is therefore density dependent, such that individuals with more elements experience greater numbers of transposition events. Particular insertions may persist in populations because of their adaptive significance, or may be co-opted for adaptation (e.g. McDonald, 1995; Lonning & Saedler, 1997; McFadden & Knowles, 1997). Considerable evidence suggests, however, that element insertions are usually harmful (Charlesworth *et al*., 1994). There are two popular models of selection against element insertions (Nuzhdin, 1999). Under the 'deleterious' model, element insertions disrupt gene function and therefore decrease organism fitness. The 'ectopic exchange' model (Montgomery *et al*., 1987; Langley *et al*., 1988), on the other hand, envisions unequal crossing over between elements at non-homologous insertion sites, resulting in harmful chromosomal rearrangements (Fig. 1: see Clegg *et al*., 1997; Caceres *et al*., 1999). For stable equilibrium element numbers under either selection model, log fitness must decline more than linearly with element number (Charlesworth, 1985; Brookfield, 1986).

Selection acting on transposable elements is expected to be weak, depending on infrequent events, and therefore difficult to observe. Indirect methods for evaluating selection models have therefore been proposed. The deleterious and ectopic exchange models make differing predictions about chromosomal localization of elements. The deleterious mutation model suggests that elements are depauperate in regions of enhanced selection, especially the X chromosome of *Drosophila* where hemizygous expression exposes heterozygous elements to selection. Ectopic exchange anticipates accumulation of elements in areas of reduced recombination, such as centromeric and telomeric regions. Empirical research supporting the plausibility and predictions of both models is available, but the interpretation of evidence is equivocal (Biemont *et al*., 1997; Charlesworth *et al*., 1997; Hoogland *et al*., 1997).

Many plants reproduce through a mixture of self-fertilization and random mating. Intuitive arguments by Charlesworth & Charlesworth (1995) suggest that differences in mating system may distinguish between the deleterious and ectopic exchange models of element regulation. Self-fertilization and other forms of inbreeding increases homozygosity.Greater homozygosity increases the strength of selection against deleterious insertions. If element number is regulated through deleterious insertion, selfing species may have fewer elements than outcrossing species. On the other hand, homozygosity may decrease the likelihood of ectopic exchange, since elements will usually have homologues to pair with. If element number is regulated through ectopic exchange, selfing species may have more elements than outcrossing species. Wright & Schoen (1999) have developed preliminary simulations of inifinite and finite populations undergoing mixed mating that support the intuitive arguments and conclusions of Charlesworth & Charlesworth. Empirical data from *Lycopersicon* species differing in mating system (Young *et al*., 1994) may support the role of ectopic exchange in element regulation (Charlesworth & Charlesworth, 1995; although see Francis *et al*., 1995). With one exception (Wright & Schoen, 1999), formal population genetic treatments of transposable element dynamics focus exclusively on random mating populations.

The objective of this study is to explore the consequences of mating system for equilibrium transposable element number, using the Ty1-*copia* group of elements as the paradigm. I develop computer simulations and analytical approximations of element transposition and selection under mixed mating. These calculations are used to determine equilibrium element number as a function of mating system, transposition rate and selection model. Simulations are also used to investigate the effects of finite population size. The results show how differing models of element regulation influence equilibrium element number under mating system variation. This provides a context for interpreting the emerging wealth of data on element number and location in outbreeding and inbreeding species.

2. Methods

(i) *Transposition, excision and selection*

Methods developed here follow element transposition, excision and selection through 'synergistic' deleterious

https://doi.org/10.1017/S0016672301005067 Published online by Cambridge University Press

insertion or 'mass action' ectopic exchange. Transposition and excision, occurring with probability u and v per element, resepectively, are assumed to occur independently for each element. If the ith individual has y_i heterozygous and z_i homozygous elements before transposition, the expected number of elements after transposition will be

$$\left.\begin{aligned} y_i' &= y_i + u(y_i + 2z_i) - v(y_i - 2z_i), \\ z_i' &= (1-2v)z_i. \end{aligned}\right\} \quad (1)$$

Transposition increases heterozygous element number in proportion to the total number of elements, $u(y_i+2z_i)$. Excision reduces heterozygous element number by vy_i and homozygous number by $\approx 2vz_i$, but each excision at a homozygous site creates a heterozygous insertion. Thus change in heterozygous elements due to excision is $-vy_i+2vz_i = -v(y_i-2z_i)$. Estimates from *Drosophila* suggest that u varies between 0·00001 and 0·001 (Nuzhdin & Mackay, 1994, 1995), with $v \approx u/10$. Estimates of transposition rates in natural plant populations are not available. Low transposition rates (e.g. $u = 0{\cdot}0001$) result in large stochastic variation in even moderate-sized populations (Charlesworth & Charlesworth, 1983; Tsitrone *et al.*, 1999). For tractability, values of $u = 0{\cdot}001$, $v = 0{\cdot}0001$ are used in most simulations. No attempt is made to include a faster than linear increase in transposition rate with element number, for which there is evidence in *Drosophila* lines with unusually high rates of movement of *copia* (Nuzhdin *et al.*, 1997).

The 'synergistic' selection model assumes that element insertions are deleterious, with effects of insertions interacting synergistically to influence individual fitness. Under this model, the fitness of an individual with y_i heterozygous and z_i homozygous elements is

$$w(y_i, z_i) = \exp\{-\alpha(hy_i + z_i) - \beta(hy_i + z_i)^2/2\}. \quad (2)$$

Heterozygous and homozygous elements influence fitness differently, as determined by the dominance coefficient h. The selection model is analogous to that used for multilocus deleterious mutation (Kondrashov, 1984, 1985; Charlesworth *et al.*, 1991; Crow, 1993), and the dominance coefficient is usually $\ll 0{\cdot}5$. Estimates of α and β are unavailable, but parameter values can be chosen to agree with empirical observation of element number in *Drosophila* and plausible parameters of selection on deleterious mutation (although see Charlesworth, 1998). Given transposition rates of $u = 0.001$ and dominance coefficient $h = 0{\cdot}2$, equilibrium element copy number in random mating populations of $\bar{y} \approx 50$ results when $\alpha \approx 8{\cdot}33 \times 10^{-4}$, $\beta = \alpha/2$.

The 'mass action' model of ectopic exchange assumes that increasing element number increases opportunity for exchange between non-homologous insertion sites. One version of fitness in this model is

$$w(y_i, z_i) = \exp\{-\omega^2(ay_i + bz_i)^2/2\}. \quad (3)$$

Fitness decreases with the square of element number, and with the rate of recombination in the genome. The dependence of fitness on element number is based on the principle that the harmful effects of ectopic exchange require participation of two elements, chosen at random and hence influencing fitness through the square of element number. Evidence for ectopic exchange is available in *Drosophila* (e.g. Montgomery *et al.*, 1991; Lim & Simmons, 1994; Caceres *et al.*, 1999) and Ac element-induced ectopic exchange is used as an experimental tool in, for instance, maize (Shalev & Levy, 1997).The mass action model of ectopic exchange assumes that chromosomal proximity of elements does not influence ectopic exchange, whereas ectopic exchange is thought to occur more commonly between elements in physical proximity (Lim & Simmons, 1994). Such localized exchange would tend to distribute elements evenly over the chromosome, and the essential feature of the mass action model (stronger selection with increasing element number) remains relevant. Equation (3) is motivated by biological realism, but relies on simplification for modelling convenience. Parameters a and b reflect the opportunity for ectopic exchange between heterozygous and homozygous insertions. The *a priori* expectation is that $b < a$, since elements at homozygous insertion sites may be less prone to non-homologous pairing. There is no direct empirical evidence for decreased ectopic exchange with increased homozygosity in plants, so various parameter values of b are investigated. The parameter ω^2 reflects the frequency with which non-homologous recombination occurs. As with α and β in (2), ω^2 can be chosen to agree with empirically observed element numbers. Models below use $\omega = 1 \times 10^{-5}$, resulting in equilibrium element number in outcrossing species of ≈ 50 when $a = 1$.

(ii) *Simulation*

Three different approaches were taken to modelling the evolutionary dynamics of transposable element evolution. The first involves detailed simulation, following chromosome insertion sites in a genome. The second approach again uses simulation, but assumes that the details of element insertion location do not matter for element evolution. Finally, analytical approximations follow mean element number in cohorts of individuals.

Computer simulations were developed to follow individual elements inserted at specific chromosome locations. The simulations model distinct individuals, each individual with several chromosomes with nu-

https://doi.org/10.1017/S0016672301005067 Published online by Cambridge University Press

merous sites for element insertion. Chromosomes of each individual are represented as strings of binary numbers, with '0' indicating no element inserted and '1' indicating element insertion at specific locations. Individuals consist of several chromosomes, in addition to summary information about the number of heterozygous and homozygous element (y_i, z_i). In the simulations reported below, individuals consist of 7 chromosomes, each experiencing an average of 1 recombination event per chromosome.

Elements are followed through stochastic transposition, mating and selection. Transposition is determined for each individual. The number of transpositions and excisions is determined by drawing a Poisson deviate from a distribution with mean defined in (1). Each new transposition is then located in the genome by drawing a uniform deviate. Insertions into sites already occupied were ignored, since such insertions are rare except when element number becomes large.

Mating starts with choice, using a uniform random deviate, of an individual to act as female parent. A second uniform deviate determines mode of reproduction (self-fertilization with probability S, random mating with probability $1-S$). Suppose the offspring is to be produced through self-fertilization. Recombination sites on each chromosome are generated in two steps. The number of recombination events is determined by drawing a Poisson deviate with mean r. The r locations of the recombination events are then placed on the chromosome using uniform deviates. Gametes are then constructed using the pattern of inheritance implied by the recombination map. Computationally, these operations are carried out using bitwise logical operations (Fraser & Burnell, 1970; Charlesworth *et al*., 1992). A second gamete is constructed using similar steps. Offspring produced through outcrossing are constructed in the same way, except that the second gamete comes from another individual chosen by drawing a uniform random deviate. The fitness of the offspring, calculated according to (2) or (3), is compared with a uniform deviate. If the fitness is less than the deviate, the offspring is rejected and the process of choosing parents and constructing gametes repeated in its entirety. This continues until N offspring are obtained.

Simulations reported below follow $N = 10000$ individuals. Populations are started with 1000 elements distributed randomly in the population. Simulations run for 5000 initial generations. A complete census is then conducted every 100 generations until generation 15000. The long series of well-spaced censuses reduces difficulties introduced by temporal correlation. The variation of census statistics for each run is typically small, less than 1% of the mean. Each parameter combination takes from one to several days of computer time (Digital Alpha 21164, 433 MHz).

(iii) '*Fast*' *simulation*

'Fast' simulations were developed to follow only the number, rather than the location, of elements in each individual. The ith individual in the simulation is characterized by the number of heterozygous and homozygous elements (y_i, z_i). Individuals are followed through stochastic transposition, mating and selection. Transposition and excision are simulated by drawing a Poisson deviate, as in (1).

Choice of parents for mating follows the procedure outlined for the full simulation. Suppose the offspring is to be produced through self-fertilization. The number of heterozygous elements in the offspring, y_i'', is determined by drawing a binomial deviate from a sample of size y_i' with probability 1/2. The number of homozygous elements, z_i'', is then chosen by drawing a binomial deviate from a sample of size $y_i' - y_i''$ with probability 1/2, and adding this to the number of homozygous elements in the parent, z_i'. Suppose instead that the offspring is produced through random mating. The number of elements in the female gamete equals z_i' plus a random deviate drawn from a sample of size y_i' with probability 1/2. The number of heterozygous elements in the outcrossed offspring, y_i'', is determined by adding numbers of elements in male and female gametes. Homozygous element number is $z_i'' = 0$. This assumes that insertions at a particular site are rare so that random mating does not result in homozygous elements. Fitnesses of offspring are calculated, and offspring accepted or rejected, as outlined in the full simulation.

Simulations reported below follow $N = 10000$ individuals. Populations are started with Poisson-distributed heterozygous elements, with mean equal to 50. Simulations run for 5000 generations to establish an approximate equilibrium.The population is then censused every 100 generations, with the final census at generation 15000. Most parameter combinations require several hours of computer time.

(iv) *Analytical approximation*

The analytical approximation follows mean heterozygous and homozygous element number in cohorts ($\bar{y}_i$, $\bar{z}_i$). Cohorts $i = 0, 1, 2, \ldots, n$ correspond to subsets of the population resulting from i generations of self-fertilization. The frequency of the ith cohort is written as p_i.

Transposition in each cohort changes heterozygous element number deterministically: $y_i' = y_i + u(y_i + 2z_i) - v(y_i - z_i)$ and $z_i' = (1-2v)z_i$. Transposition does not change cohort frequency, so $p_i' = p_i$.

Reproduction occurs through self-fertilization, probability S, or random mating, probability $1-S$. The transition in element number associated with self-fertilization is $\bar{y}_{i+1}'' = \bar{y}_i'/2$ and $\bar{z}_{i+1}'' = \bar{z}_i' + \bar{y}_i'/4$. The

https://doi.org/10.1017/S0016672301005067 Published online by Cambridge University Press

frequency of the $(i+1)$th cohort after self-fertilization is $p''_{i+1} = Sp'_i$. The mean element number in the nth (i.e. final) cohort is calculated as the average of the nth and $(n-1)$th cohorts, $p''_n = S(p'_n + p'_{n-1})$. Random mating determines element number in cohort 0. Random mating is assumed to result only in heterozygous elements (i.e. selection is sufficiently strong that elements are individually rare). The number of heterozygous elements is calculated as the average number of heterozygous elements in gametes contributed by each cohort. Thus $\bar{y}''_0 = 2\sum_{i=0}^{n} p'_i(\bar{y}'_i/2 + \bar{z}'_i)$ and $\bar{z}''_0 = 0$. The frequency of the 0th cohort after random mating is $p''_0 = (1-S)$.

Change in element number due to selection is based on a quantitative genetic approximation. The change in the mean of trait x due to selection is $\Delta\bar{x} \approx \sigma_x^2 \, \partial \ln \bar{w} / \partial \ln \bar{x}$ (Lande, 1976; Falconer & Mackay, 1996). The variance of x is σ_x^2, and it is assumed that trait x is uncorrelated with other traits. Selection is applied within each cohort, with cohort mean fitness, $\bar{w}_i$, changing the frequency of the cohort, $p^*_i = p''_i \bar{w}_i / \bar{w}$. Cohort mean fitness is approximated as the fitness at the cohort average element number, $\bar{w}_i \approx w(\bar{y}_i, \bar{z}_i)$. Several authors discuss conditions when this approximation is accurate (e.g. Abrams *et al.*, 1993). Specific fitness functions are given above for synergistic (equation 2) and mass action (equation 3) models of selection. Within each cohort, the distribution of heterozygous and homozygous elements is assumed to follow a bivariate Poisson distribution. The variance of heterozygous and homozygous element numbers is then $(\bar{y}''_i, \bar{z}''_i)$. It is assumed that there is no correlation between heterozygous and homozygous elements between individuals within each cohort. Justification for these assumptions is threefold. Theoretical arguments suggest that element number should be approximately Poisson-distributed in random mating populations (Langley *et al.*, 1983). Quantitative genetic studies of selection in mixed mating species have successfully applied a cohort approach (Kelly, 1999*a*, *b*), partly because identity disequilibrium (cf. Lande *et al.*, 1994) generated by self-fertilization is likely to exceed greatly that generated by selection. Finally, tedious calculation following both mean and variance of element numbers within cohorts (using equations modified from Charlesworth *et al.*, 1991) lead to results very similar to those obtained here (M. T. Morgan, unpublished.).

The foregoing outlines how transposition, mating and selection influence element number over a single generation. Roots of the resulting non-linear, complicated set of equations can be solved explicitly. Here, however, equations were iterated until proportional change in average element number between generations was less than 10^{-6}. The number of cohorts followed was set to 20, although variation in cohort number greater than ≈ 10 has little effect on results obtained. The convergence criterion is usually satisfied after a few seconds of computer time.

3. Results

(i) *Synergistic deleterious insertion*

Fig. 2 shows equilibrium element number under synergistic selection. When net transposition, $u-v$, equals 0·001, greater self-fertilization reduces equilibrium element number. Reduction in element number makes intuitive sense, because self-fertilization increases the number of homozygous relative to heterozygous elements, and hence the strength of selection against elements (Kimura & Maruyama, 1966). In addition, self-fertilization increases the variance in element number between individuals, allowing greater opportunity for selection (Burt & Trivers, 1998). Complete self-fertilization reduces element number more than 50-fold. This dramatic reduction reflects the feedback effect of reduced element number on rate of transposition: stronger selection reduces element number, effectively decreasing the number of elements undergoing transposition each iteration. Analytical approximation agrees closely with both the full and fast simulations when selfing rate is less than 0·8 (Fig. 2, open symbols and lines). Discrepancies at high selfing rate occur, when all elements are lost from the simulations. Apparently, equilibrium element number decreases sufficiently that genetic drift can result in fixation of element-free genotypes, even in large populations. The effects of population size are discussed in greater detail below.

Qualitatively different patterns of element number occur when transposition rate becomes large (e.g. $u =$ 0·01 in Fig. 2). Element number remains high and approximately constant until intermediate selfing ($S \approx 0.62$), and then decreases rapidly. These analytical results are supported by fast simulations. Unfortunately, the large number of elements involved precludes full simulation of this effect. A similar threshold is observed in models of deleterious mutation, provided that the genomic mutation rate is large relative to the dominance coefficient (Lande *et al.*, 1994). The heuristic explanation for this pattern is that weak selection allows heterozygous elements to accumulate to high levels. Selfed progeny then have many homozygous loci and extremely low fitness. The post-selection population consists almost exclusively of outcrossed individuals. Under these circumstances, homozygosity associated with self-fertilization is ineffective at reducing average element number.

Fig. 3 investigates the threshold effect in more detail, using results from the analytical model. An abrupt threshold like that in Fig. 2 is observed for all net rates of transposition. There is a direct relationship

https://doi.org/10.1017/S0016672301005067 Published online by Cambridge University Press

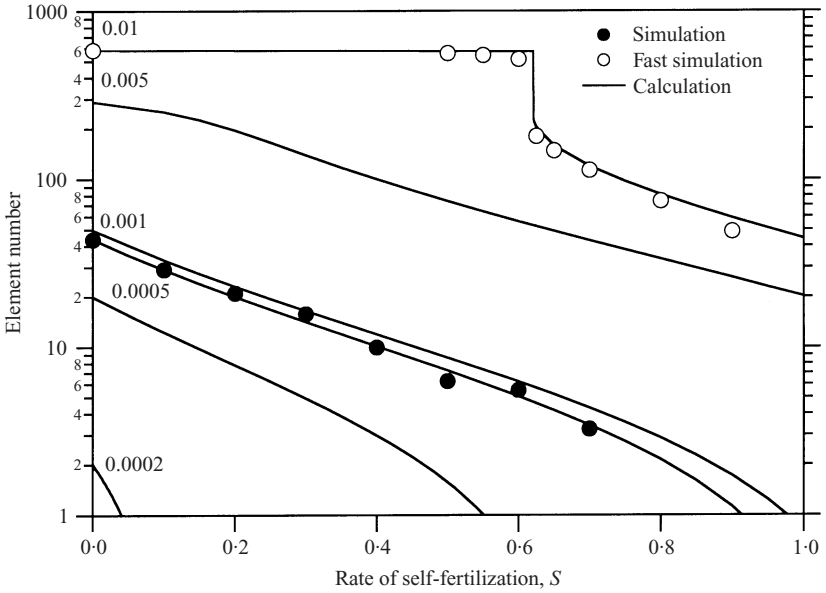

Fig. 2. Equilibrium element number under synergistic selection ($\alpha = 8{\cdot}33 \times 10^{-4}$, $\beta = \alpha/2$), with net rates of transposition ($u-v$; $u = 10v$) varying between 0·0002 and 0·01.

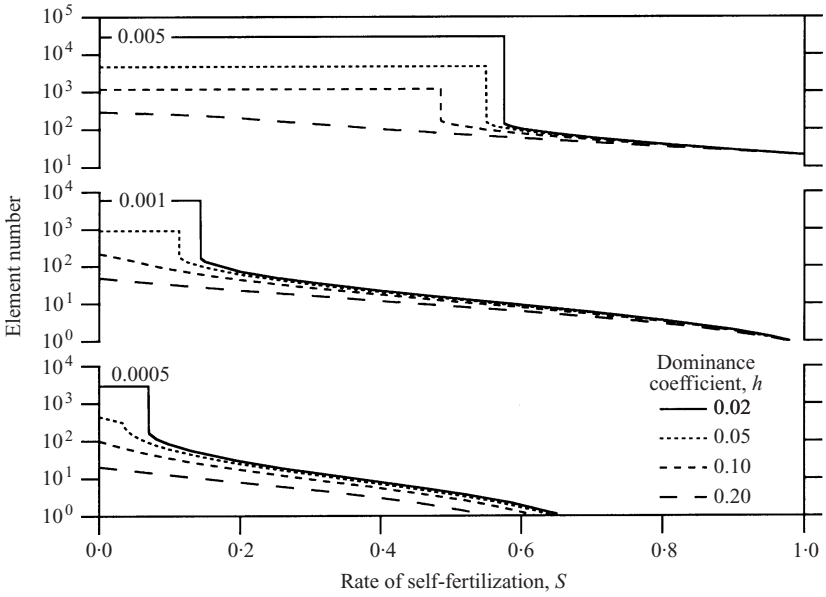

Fig. 3. Equilibrium element number under synergistic selection ($\alpha = 8{\cdot}33 \times 10^{-4}$, $\beta = \alpha/2$), with dominance h between 0·02 and 0·2, and net transposition $u-v$; $u = 10v$ between 0·0005 and 0·005.

between transposition rate and dominance coefficient, required for the threshold to appear. For instance, when $u-v = 0{\cdot}01$ a dominance coefficient of $h = 0{\cdot}2$ generates a threshold (Fig. 2), while $u-v = 0{\cdot}0005$ exhibits a threshold only at $h \leqslant 0{\cdot}05$ (Fig. 3). Increasing the transposition rate or decreasing the dominance coefficient increases the selfing rate at which the threshold occurs (Fig. 3).

(ii) *Mass action ectopic exchange*

Fig. 4 shows equilibrium element number under the mass action model of ectopic exchange.When element insertions are additive at individual sites ($a = 2b$), self-fertilizing populations have half as many elements as outcrossed populations. This makes intuitive sense, as each selective elimination removes two elements in the selfer rather than one in the outcrosser (Kimura & Maruyama, 1966). It is more biologically plausible to assume that homozygous elements are less prone to ectopic exchange than heterozygous elements ($a \geqslant b$). Self-fertilization then results in an increase in element number. The effect tends to be gradual over most of the range of selfing rates. Decreasing effects of homozygous insertions delays increase in element number until a higher selfing rate, but the increase becomes more dramatic. Simulation and analytical calculations are in agreement when selfing rate is less

https://doi.org/10.1017/S0016672301005067 Published online by Cambridge University Press

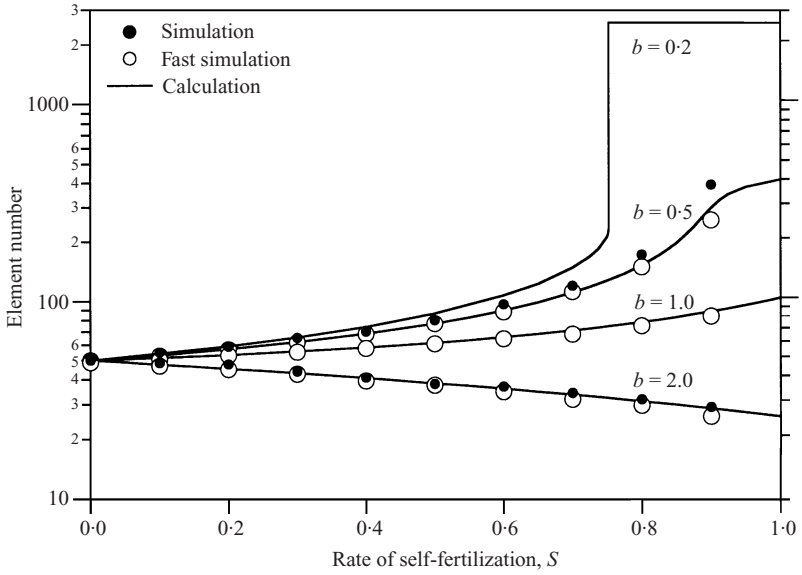

Fig. 4. Equilibrium element number under mass action ectopic exchange, with selection parameters $a = 1$, $\omega = 1 \times 10^{-5}$ and net transposition $u - v = 0{\cdot}001$.

than $\approx 0{\cdot}8$. At high selfing rates and when $a \gg b$, element number increases to the computational limits of simulation. Stochastic effects at high selfing rate are discussed further below.

With small values of b, there is a threshold complementary to that observed under synergistic selection with low dominance coefficients. The reason for this is comparable to the selective interference under synergistic selection, but with self-fertilized (homozygous) individuals having high fitness compared with outcrossed (heterozygous) individuals. Additional results show that the occurrence and location of the threshold is influenced by transposition rate and relative selective effects on heterozygous and homozygous elements: the threshold occurs at higher selfing rate when b decreases or net transposition increases.

(iii) *Mean fitness and modifier selection*

Fig. 5 depicts the relationship between population mean fitness and rate of net transposition under different mating systems. For all models of selection, greater net transposition decreases mean fitness. Mating system does not affect the decrease in mean fitness under mass action ectopic exchange when $b = 1$ (lower panel; $b = 1$ corresponds to equal occurrence of ectopic exchange at heterozygous and homozygous sites). With synergistic selection (upper panel) and with ectopic exchange occuring more frequently in heterozygotes than homozygotes (lower panel, $b = 0{\cdot}5$), mean fitness decreases much more rapidly in outcrossing than self-fertilizing populations.

The slope of the relationship between the logarithm of mean fitness and rate of net transposition, $d \ln \bar{w}/d(u-v)$, is the selection gradient acting on individual net transposition. Under synergistic selection models, selection for reduced net transposition is stronger in outcrossing populations than in self-fertilizing populations. With mass action ectopic exchange, fitness parameters resulting in greater element number (i.e. $b < 1$, Fig. 4) enhance selection against net transposition much more strongly in self-fertilizing than in outcrossing populations. Evolutionary modification of the net transposition rate therefore obscures the mating system differences in element number outlined above.

(iv) *Population size*

Fast simulations show that the usual consequence of finite size under synergistic selection is element loss (Table 1). Simulations varying the rate of excision u (results not shown) indicate that the loss of elements occurs primarily through stochastic fixation of element-free sites, rather than element excision. Element loss occurs more than half the time in very small ($N = 50$) outcrossing populations. The threshold for element persistence when $S > 0{\cdot}5$ can be large; populations selfing at $S = 0{\cdot}9$ must be larger than $N = 10000$ individuals for elements to persist in at least half the replicate populations. This probably reflects the low equilibrium frequency expected under self-fertilization, with small fluctuations in average element number sufficient to result in element loss.

Table 2 documents unusual behaviour in small populations under synergistic selection with complete self-fertilization ($S = 1{\cdot}0$). While elements are lost from 68–92 % of replicates (Table 1), elements increase to substantial frequency in the remaining simulations. Long-term average element number in those simulations where elements persist decrease with increasing

https://doi.org/10.1017/S0016672301005067 Published online by Cambridge University Press

Table 1. *Influence of population size and selfing rate on element fate under synergistic selection*

Population size, N	Rate of self-fertilization, S					
	0·0	0·5	0·8	0·9	0·95	1·0
50	58/0	100/0	100/0	100/0	100/0	*74/0*
100	12/0	100/0	100/0	100/0	100/0	*68/0*
200	2/0	98/0	100/0	100/0	100/0	*86/0*
500	0/0	54/0	100/0	100/0	100/0	*92/0*
1000	−/−	22/0	96/0	100/0	100/0	100/0
2000	−/−	2/0	78/0	96/0	96/0	100/0
5000	−/−	0/0	32/0	78/0	4/0	100/0
10000	−/−	−/−	6/0	64/0	88/0	100/0
20000	−/−	−/−	0/0	32/0	80/0	100/0

Entries have the form A/B, where A is the percentage of simulations in which elements were lost and B is the percentage of simulations where average element number exceeds 1000. Italicized results are discussed in the text and Table 2.

Table 2. *Unusual equilibrium numbers of heterozygous, $\bar{y}$, and homozygous, $\bar{z}$, elements in small populations at high selfing rates ($S = 1{\cdot}0$) under synergistic selection*

Population size, N	$\bar{y}$	$\bar{z}$
50	0·73	150
100	0·34	70·9
200	0·15	32·3
500	0·035	7·25

population size (Table 2). Average element numbers appear to represent an approximate equilibrium, since populations persist for over 150000 iterations.

Finite population size under ectopic exchange results in loss of elements in outcrossing populations (Table 3). Drift presumably overcomes weak transposition in finite outcrossing populations, as also occurs with synergistic selection. With increasing self-fertilization, the deterministic equilibrium element number increases. This accounts for the reduced likelihood of drift eliminating elements at $S \geqslant 0{\cdot}5$.

At higher selfing rates under ectopic exchange, element number sometimes increases explosively (Table 3; $S \geqslant 0{\cdot}8$). When $S > 0{\cdot}95$ only very large populations (> 10000) are able to moderate element number.The dynamics of element number increase in these situations follow a characteristic pattern. Populations reach a quasi-equilibrium, with mean fitness and heterozygote and homozygote number fluctuating around deterministic equilibrium levels. Over a short period of time (several hundred iterations), element number shifts so that on average there are many homozygous and few heterozygous elements. The underlying reason for the explosive increase in element number is probably related to selective interference (Lande *et al.*, 1994), evidenced in deterministic calculations by the abrupt increase in

Table 3. *Influence of population size and selfing rate on element fate under mass action ectopic exchange.*

Population size, N	Rate of self-fertilization, S					
	0·0	0·5	0·8	0·9	0·95	1·0
50	46/0	34/0	4/68	0/100	0/100	0/100
100	12/0	0/0	0/16	0/100	0/100	0/100
200	0/0	−/−	0/0	0/100	0/100	0/100
500	−/−	−/−	−/−	0/88	0/100	0/100
1000	−/−	−/−	−/−	0/14	0/100	0/100
2000	−/−	−/−	−/−	0/0	0/100	0/100
5000	−/−	−/−	−/−	−/−	0/92	0/100
10000	−/−	−/−	−/−	−/−	0/36	0/100
20000	−/−	−/−	−/−	−/−	0/2	0/100

Entries have the form A/B, where A is the percentage of simulations in which elements were lost and B is the percentage of simulations where average element number exceeds 1000.

https://doi.org/10.1017/S0016672301005067 Published online by Cambridge University Press

element number at $S \approx 0.8$ in Fig. 4. The high number of homozygous elements means that outcrossed progeny have very low fitness compared with selfed progeny, so the population consists of highly inbred lineages. Element number continues to increase rapidly once inbred lineages become established. The rapid increase reflects the positive feedback of transposition, with greater element number facilitating increased numbers of new transpositions. Element number appears to increase without limit ($\bar{n} > 4000$) in smaller highly selfing populations.

4. Discussion

The models of element regulation investigated here predict different patterns of element number in self-fertilizing compared with outcrossing species. When element number is regulated by the deleterious effects of insertion (synergistic selection), increasing self-fertilization decreases element number. Weak dominance results in a threshold effect where element number only declines above an intermediate selfing rate. In many predominantly selfing populations, genetic drift is likely to result in loss of all elements. Element number may, however, increase in highly selfing populations ($S > 0{\cdot}95$). When element number is regulated through ectopic exchange, and when homozygous elements are less likely to exhibit ectopic exchange than heterozygous elements ($b \leqslant 1$, Fig. 1), self-fertilization increases equilibrium element number. The increase is usually gradual, but can be abrupt in small populations or at high selfing rates and when homozygous insertions are much less likely than heterozygous insertions to participate in ectopic exchange. Accumulation of element number under ectopic exchange in self-fertilizing species is accompanied by increasingly strong selection for decreased net transposition (Fig. 5).

The following summarizes existing empirical studies on element number and localization.The summary provides some support for the importance of deleterious insertion, rather than ectopic exchange, in regulating element number.

(i) *Element number*

Table 4 summarizes empirical estimates of Ty1-*copia* element number in relation to mating system. The Ty1-*copia* element family is the best characterized of plant transposable elements. Most studies estimating element number investigate economically important or model species. Element number is usually assessed using heterologous probes or primers. This results in qualitative estimates, especially when element number exceeds ≈ 100. Mating system characterization is from Tables 1.4 and 1.5 of Frankel & Galun (1977). Outbreeding taxa typically have many elements ($\geqslant 10\,000$). Inbreeding taxa frequently have fewer elements ($\leqslant 1000$), with prominent exceptions (*Hordeum vulgare*, *Oryza australiensis*, *Vicia faba*). This pattern is more consistent with deleterious insertion, rather than ectopic exchange, but statistical support for the pattern is weak. Non-parametric analysis, excluding species with only qualitative element number estimates, results in marginally significant differences (Kolmogorov–Smirnov test, $p \approx 0{\cdot}044$); differences are not significant using a parametric *t*-test. The statistical analysis is extremely

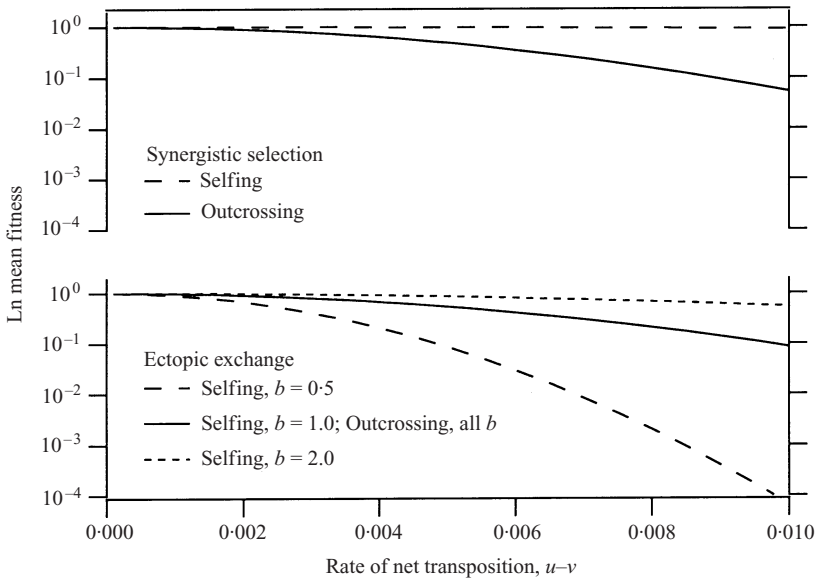

Fig. 5. Relationship between mean fitness (logarithmic scale) and net rate of transposition ($u - v$; $u = 10v$) under synergistic selection ($\alpha = 8{\cdot}33 \times 10^{-4}$, $\beta = \alpha/2$, $h = 0{\cdot}2$) and mass action ectopic exchange ($a = 1$, $\omega = 1 \times 10^{-5}$, values of b as indicated). 'Selfing' values correspond to $S = 0{\cdot}99$, 'outcrossing' values to $S = 0{\cdot}01$. Mean fitness does not depend on mating system when $b = 1$.

https://doi.org/10.1017/S0016672301005067 Published online by Cambridge University Press

Table 4. *Empirical relationship between mating system and Ty1*-copia-*like element number*

Species	Number	Reference
Inbreeding		
Arabidopsis thaliana	200	Brandes *et al.*, 1997; < 3 per family, 28 families reported by Voytas & Ausubel, 1988; Voytas *et al.*, 1990; Konieczny *et al.*, 1991; Wright *et al.*, 1996; 'several' copies from 100 families, Kapitonov & Jurka, 1999; 14% of genomic sequence, Ty1-*copia* fraction not specified, The Arabidopsis Genome Initiative, 2000
Cicer arietinum	600	Sant *et al.*, 2000
?*C. reticulatum*	10	Sant *et al.*, 2000
Citrus spp.	Multiple	Asins *et al.*, 1999
Glycine max	500–800	Laten & Morris, 1993
Gossypium spp.	100s–1000s	VanderWiel *et al.*, 1993
Hordeum vulgare	≈ 70000	Suoniemi *et al.*, 1996; Vicient *et al.*, 1999
?*Lycopersicon chilens*	2200	Yanez *et al.*, 1998
L. esculentum	200	Unpublished, in Kumar *et al.*, 1997
Nicotiana tabacum	> 100	Grandbastien *et al.*, 1989; Pouteau *et al.*, 1991
	> 30	*Tto1*, Hirochika, 1993
Oryza spp.	1000	Hirochika *et al.*, 1992
	1–5	*ToS17*, Hirochika *et al.*, 1992
O. sativa	Moderate/high	Wang *et al.*, 1997; < 1·5% of genome, Mao *et al.*, 2000
? *O. australiensis*	Extraordinary	Nakajima *et al.*, 1996; Noma *et al.*, 1997
Phaseolus vulgaris	40	Garber *et al.*, 1999
Pisum sativum	50	Grandbastien, 1992
Solanum tuberosum	> 360	Camirand *et al.*, 1990; Flavell *et al.*, 1992*b*
Triticum aestivum	200	Harberd *et al.*, 1987; Moore *et al.*, 1991
Vicia faba	1000000	Pearce *et al.*, 1996*a*
?*V. melanops*	1000	Pearce *et al.*, 1996*a*
V. sativa	5000	Pearce *et al.*, 1996*a*
Vitus spp.	> 20	Verries *et al.*, 2000
Outbreeding		
Allium cepa	10000	Pearce *et al.*, 1996b
Beta vulgaris	High	Schmidt *et al.*, 1995
Elaeis guineensis	'Relatively low'	Castilho *et al.*, 2000
Prunus spp.	1–3	Asins *et al.*, 1999
Pinus spp.	High	Kamm *et al.*, 1996
Secale cereale	100000	Pearce *et al.*, 1997
Zea maize	> 30000	*Opie 1*, SanMiguel *et al.*, 1996
	> 10000	PREM-2, Turcich *et al.*, 1996
	1–5	*Bs1*, Jin & Bennetzen, 1989

Mating system characterized by Frankel & Galun (1977). The mating system of taxa preceded by a question mark (?) is uncertain.

sensitive to the necessarily approximate counts from some species, and to inclusion of extreme counts in either mating system category.

There are several obvious limitations to the broad-scale comparative study in Table 4. Characterization of mating system is usually qualitative and historically recent, rather than a quantitative estimate of the evolutionarily relevant (i.e. on the timescale during which element number evolves) rate of self-fertilization. For instance, the large showy flowers and natural history of *Gossypium* suggest substantial outcrossing (Fryxell, 1979), whereas cultivated cotton is characterized as inbreeding (Frankel & Galun, 1977). Exceptional element numbers in *H. vulgare*, *O. australiensis* and *V. faba* might reflect high selfing or reductions in population size during domestication (see Table 1). The largely qualitative assessment of element number may also obscure an underlying pattern. Most studies identify short homologous portions of single genes. This does not allow distinction between intact, active elements and fragmented or otherwise inactive elements. In chromosome 2 of *Arabidopsis thaliana*, for instance, over 50% of transposon sequences are pseudogenes (Lin *et al.*, 1999). Genome size varies amongst species, potentially altering the number of suitable insertion sites (Kumar *et al.*, 1997). In particular, many of the species in Table 4 are polyploid (VanderWiel *et al.*, 1993). Consequences of polyploidy are not investigated in the simulations reported here. Finally, domestication

https://doi.org/10.1017/S0016672301005067 Published online by Cambridge University Press

Table 5. *Chromosomal location of Ty1*-copia *elements, determined using* in situ *hybridization*

Species	Distribution
Inbreeding	
Arabidopsis thaliana	Pericentromeric heterochromatin (Brandes *et al*., 1997)
Avena spp.	Dispersed. −, Centromeric and nucleolar organizer regions (Katsiotis *et al*., 1996)
Cicer arietinum	Pericentromeric heterochromatin (Brandes *et al*., 1997)
Hordeum vulgare	Dispersed. −, Pericentromeric regions, telomeres and nucleolar organizer regions (Suoniemi *et al*., 1996)
Oryza sativa	Dispersed (Mao *et al*., 2000)
Sorghum vulgare	Dispersed (Miller *et al*., 1998)
Vicia faba, *V. narbonensis*	Dispersed. −, Centromeric, nucleolar organizing, and tandem repeat regions. +, Local (Brandes *et al*., 1997)
Outbreeding	
Allium cepa	Dispersed. +, Terminal heterochromatin (Pearce *et al*., 1996*b*)
Brassica campestris, *B. oleracea*	Dispersed. −, Centromeric, nucleolar organizing and tandem repeat regions. +, Local (Brandes *et al*., 1997)
Beta vulgaris	Dispersed. −, Local (Schmidt *et al*., 1995)
Pennisetum glaucum	Dispersed. −, Centromeric, nucleolar organizing and tandem repeat regions. +, Local (Brandes *et al*., 1997)
Pinus elliottii	Dispersed. −, Centromeric; 18S, 5·8S, 25S rRNA (Kamm *et al*., 1996)
Secale cereale	Dispersed. −, Centromeric regions and terminal heterochromatin (Pearce *et al*., 1997)
Zea mays	Dispersed (Miller *et al*., 1998)

Mating system characterized by Frankel and Galun (1977).
+/− indicates regions where element numbers are enriched or reduced.

may introduce features, such as population bottlenecks, that influence element numbers in unexpected ways.

The theory developed here provides motivation for surveying element numbers across related taxa differing in mating system. Charlesworth & Charlesworth (1995) suggested that element number in *Lycopersicon* spp. (Young *et al*., 1994) is consistent with ectopic exchange. This is based on the presence of fewer Lyt*1*-hybridizing sequences in green-fruited (putatively outcrossing) species compared with red-fruited (putatively self-fertilizing) species. Green-fruited species appear to contain 2–8 copies, compared with 40–50 in red-fruited species. This is consistent with much less ectopic exchange between homozygous compared with heterozygous elements (e.g. $b < 0{\cdot}5$ in Fig. 4). Rigorous conclusions are undermined by uncertainty surrounding the evolutionarily relevant selfing rates of these species (Francis *et al*., 1995). Candidates for additional study identified in Table 4 include *Vicia* spp., *Pinus* spp. and *Oryza* spp. The contrast between inbreeding *Arabidopsis* and outbreeding *Arabis* is also relevant.

(ii) *Chromosomal localization*

Several authors use chromosomal localization to distinguish between theories of element regulation (Hoogland & Biemont, 1996; Biemont *et al*., 1997; Charlesworth *et al*., 1997; Duret *et al*., 2000). In random mating populations under ectopic exchange, element abundance decreases with greater rates of recombination (Langley *et al*., 1988; Charlesworth & Langley, 1989), including on the X chromosome where recombination in females is possible (Langley *et al*., 1988). This assumes that ectopic and meiotic exchange have similar densities, since patterns of ectopic exchange have not been directly documented. Under deleterious insertion, elements are underrepresented in regions where selection is strong (i.e., regions of high gene density). Both theories anticipate lower element abundance in regions of enhanced recombination due to the Hill–Robertson effect (Duret *et al*., 2000). These predictions change under self-fertilization. Homozygosity reduces opportunity for ectopic exchange (provided $b < 1$), regardless of the rate of recombination in surrounding chromosomal regions. Elements may therefore be dispersed throughout chromosomes of inbreeding species when element regulation occurs through ectopic exchange. In contrast, homozygosity associated with inbreeding may amplify the deleterious consequences of element insertion into functional genes. Elements may therefore be restricted to genomic regions of inbreeding species with few active genes when element regulation occurs through deleterious insertion. Self-fertilization effectively reduces recombination, greatly reducing the importance of the Hill–Robertson effect in influencing chromosomal localization.

Chromosomal location of Ty1-*copia* elements in plants has been studied primarily using *in situ* hybridization (Table 5). Ty1-*copia* elements are

https://doi.org/10.1017/S0016672301005067 Published online by Cambridge University Press

dispersed throughout chromosomes in all species investigated, except for *A. thaliana* and the inbreeding *Cicer arietinum*. Elements are often absent or at reduced frequency in centromeric, pericentromeric or nucleolar organizing regions, and occasionally enriched in pericentromeric or terminal heterochromatin. There is no obvious patterning of element localization with mating system, with the exception of restricted distribution in inbreeding *A. thaliana* and *C. arietinum*.

Genome sequencing is providing an exquisite look at element localization in the inbreeding *A. thaliana*. Transposable elements are strongly clustered in the pericentromeric region of each chromosome and in regions of low gene expression (The Arabidopsis Genome Initiative, 2000). Clustering of elements near the centromere has been taken as evidence for the importance of ectopic exchange, because recombination is often thought to be reduced in this region (Langley *et al*., 1988; Charlesworth & Langley, 1989). Although recombination is severely reduced at the centromere, recombination in the pericentromeric regions is suppressed only modestly in *A. thaliana* (Copenhaver *et al*., 1999; Lin *et al*., 1999). In addition, as argued above, restricted recombination is less relevant for ectopic exchange (when $b < 1$) in a highly inbreeding species such as *A. thaliana*. The abundance of transposable elements in regions of low gene density may, therefore, represent preferential persistence of insertions in non-coding sequence. Duret *et al*. (2000) have recently shown that transposons but not retrotransposons occur in regions of high recombination in self-feritlizing *Caenorhabditis elegans*. These authors argue that this pattern is inconsistent with both ecotpic exchange and deleterious insertion models of element regulation, because the Hill–Robertson effect anticipates element accumulation in areas of restricted recombination. As outlined above, though, self-fertilization eliminates much of the opportunity for chromosomal localization via the Hill–Robertson effect.

(iii) *Summary and future directions*

Computer simulations support previous qualitative predictions about how mating system influences element number under different modes of regulation. Broadly speaking, greater self-fertilization decreases element number under deleterious insertion and increases element number with regulation through ectopic exchange. Empirical data, though far from satisfactory, are more consistent with the deleterious effects of insertion than with ectopic exchange. Specifically, element number in inbreeding species is generally reduced compared with outbreeding species. Elements are generally dispersed throughout chromosomes, rather than being restricted to areas of restricted recombination as anticipated by regulation through ectopic exchange. In inbreeding *A. thaliana*, element localization to the pericentromeric region probably reflects the decreased density of coding genes in this region. Unfortunately, this is consistent with both deleterious insertion and ectopic exchange. The simplistic nature of the population genetic models presented here, and the correspondingly crude characterization of element number and localization, emphasize the opportunity for additional theoretical and empirical research into factors regulating element number.

This work received partial support from the US National Science Foundation grant DEB-9974017.

References

Abrams, P. A., Harada, Y. & Matsuda, H. (1993). On the relationship between quantitative genetic and ESS models. *Evolution* **47**, 982–985.

Asins, M. J., Monforte, A. J., Mestre, P. F. & Carbonell, E. A. (1999). *Citrus* and *Prunus copia*-like retrotransposons. *Theoretical and Applied Genetics* **99**, 503–510.

Bennetzen, J. L. (1996). The contributions of retroelements to plant genome organization, function and evolution. *Trends in Microbiology* **4**, 347–353.

Berg, D. E. & Howe, M. M. (1989). *Mobile DNA*. Washington, DC: American Society for Microbiology.

Bhattacharyya, M. K., Smith, A. M., Ellis, T. H., Hedley, C. & Martin, C. (1990). The wrinkled-seed character of pea described by Mendel is caused by a transposon-like insertion in a gene encoding starch-branching enzyme. *Cell* **60**, 115–122.

Biemont, C., Tsitrone, A. & Hoogland, C. (1997). Transposable element distribution in *Drosophila*. *Genetics* **147**, 1997–1999.

Brandes, A., Heslop Harrison, J. S., Kamm, A., Kubis, S., Doudrick, R. L. & Schmidt, T. (1997). Comparative analysis of the chromosomal and genomic organization of Ty1-*copia*-like retrotransposons in pteridophytes, gymnosperms and angiosperms. *Plant Molecular Biology* **33**, 11–21.

Brookfield, J. F. Y. (1986). A model of DNA sequence evolution within transposable elements. *Genetics* **112**, 393–407.

Burt, A. & Trivers, R. (1998). Selfish DNA and breeding system in flowering plants. *Proceedings of the Royal Society of London, Series B* **265**, 141–146.

Caceres, 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.

Camirand, A., St-Pierre, B., Marineau, C. & Brisson, N. (1990). Occurrence of a *copia*-like transposable element in one of the introns of the potato starch phosphorylase gene. *Molecular and General Genetics* **224**, 33–39.

Capy, P., Gasperi, G., Biemont, C. & Bazin, C. (2000). Stress and transposable elements: co-evolution or useful parasites? *Heredity* **85**, 101–106.

Castilho, A., Vershinin, A. & Heslop Harrison, J. S. (2000). Repetitive DNA and the chromosomes in the genome of oil palm (*Elaeis guineensis*). *Annals of Botany* **85**, 837–844.

Charlesworth, B. (1985). The population genetics of

transposable elements. In *Population Genetics and Molecular Evolution* (ed. T. Ohta & K. Aoki), pp. 213–232. Tokyo: Japan Science Society Press.
Charlesworth, B. (1987). The heritability of fitness. In *Sexual Selection: Testing the Alternatives* (ed. J. W. Bradbury & M. B. Andersson), pp. 21–40. New York: Wiley.
Charlesworth, B. (1998). The effect of synergistic epistasis on the inbreeding load. *Genetical Research* **71**, 85–89.
Charlesworth, B. & Charlesworth, D. (1983). The population dynamics of transposable elements. *Genetical Research* **42**, 1–27.
Charlesworth, B. & Langley, C. H. (1989). The population genetics of *Drosophila* transposable elements. *Annual Reviews of Genetics* **23**, 251–287.
Charlesworth, B., Morgan, M. T. & Charlesworth, D. (1991). Multilocus models of inbreeding depression with synergistic selection and partial self-fertilization. *Genetical Research* **57**, 177–194.
Charlesworth, B., Sniegowski, P. & Stephan, W. (1994). The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* **371**, 215–220.
Charlesworth, B., Langley, C. H. & Sniegowski, P. D. (1997). Transposable element distributions in *Drosophila*. *Genetics* **147**, 1993–1995.
Charlesworth, D. & Charlesworth, B. (1995). Transposable elements in inbreeding and outbreeding populations. *Genetics* **140**, 415–417.
Charlesworth, D., Morgan, M. T. & Charlesworth, B. (1992). The effect of linkage and population size on inbreeding depression due to mutational load. *Genetical Research* **59**, 49–61.
Clegg, M. T., Cummings, M. P. & Durbin, M. L. (1997). The evolution of plant nuclear genes. *Proceedings of the National Academy of Sciences of the USA* **94**, 7791–7798.
Copenhaver, G. P., Nickel, K., Kuromori, T., Benito, M. I., Kaul, S., Lin, X., Bevan, M., Murphy, G., Harris, B., Parnell, L. D., McCombie, W. R., Martienssen, R. A., Marra, M. & Preuss, D. (1999). Genetic definition and sequence analysis of *Arabidopsis* centromeres. *Science* **286**, 2468–2474.
Crow, J. F. (1993). Mutation, mean fitness, and genetic load. *Oxford Surveys in Evolutionary Biology* **9**, 3–42.
Duret, L., Marais, G. & Biemont, C. (2000). Transposons but not retrotransposons are located preferentially in regions of high recombination rate in *Caenorhabditis elegans*. *Genetics* **156**, 1661–1463.
Falconer, D. S. & Mackay, T. F. (1996). *Introduction to Quantitative Genetics*, 4th edn. Harlow, Essex: Longman.
Flavell, A. J., Dunbar, E., Anderson, R., Pearce, S. R. & Hartley, R. (1992*a*). Ty1-*copia* group retrotransposons are ubiquitous and heterogeneous in higher plants. *Nucleic Acids Research* **20**, 3639–3644.
Flavell, A. J., Smith, D. B. & Kumar, A. (1992*b*). Extreme heterogeneity of Ty1-*copia* group retrotransposons in plants. *Molecular and General Genetics* **231**, 233–242.
Francis, D. M., Aitken, K. S. & St Clair, D. A. (1995). The relationship between mating system and *Lyt* element abundance in *Lycopersicon* and *Solanum*. *Genetics* **140**, 418–419.
Frankel, R. & Galun, E. (1977). *Pollination Mechanisms, Reproduction and Plant Breeding*. Berlin: Springer.
Fraser, A. S. & Burnell, D. (1970). *Computer Models in Genetics*. New York: McGraw-Hill.
Fryxell, P. A. (1979). *The Natural History of the Cotton Tribe*. College Station: Texas A & M University Press.
Garber, K., Bilic, I., Pusch, O., Tohme, J., Bachmair, A., Schweizer, D. & Jantsch, V. (1999). The Tpv2 family of retrotransposons of *Phaseolus vulgaris*: structure, integration characteristics, and use for genotype classification. *Plant Molecular Biology* **39**, 797–807.
Grandbastien, M. A. (1992). Retroelements in higher plants. *Trends in Genetics* **8**, 103–108.
Grandbastien, M. A., Spielmann, A. & Caboche, M. (1989). Tnt1, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. *Nature* **337**, 376–380.
Harberd, N. P., Flavell, R. B. & Thompson, R. D. (1987). Identification of a transposon-like insertion in a Glu-1 allele of wheat. *MGG: Molecular and General Genetics* **209**, 326–332.
Hirochika, H. (1993). Activation of tobacco retrotransposons during tissue culture. *EMBO Journal* **12**, 2521–2528.
Hirochika, H. & Hirochika, R. (1993). Ty1-*copia* group retrotransposons as ubiquitous components of plant genomes. *Japanese Journal of Genetics* **68**, 35–46.
Hirochika, H., Fukuchi, A. & Kikuchi, F. (1992). Retrotransposon families in rice. *Molecular and General Genetics* **233**, 209–216.
Hoogland, C. & Biemont, C. (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.
Hoogland, C., Vieira, C. & Biemont, C. (1997). Chromosomal distribution of the 412 retrotransposon in natural populations of *Drosophila simulans*. *Heredity* **79**, 128–134.
Jin, Y. K. & Bennetzen, J. L. (1989). Structure and coding properties of Bs1, a maize retrovirus-like transposon. *Proceedings of the National Academy of Sciences of the USA* **86**, 6235–6239.
Kamm, A., Doudrick, R. L., Heslop-Harrison, J. S. & Schmidt, T. (1996). The genomic and physical organization of Ty1-*copia*-like sequences as a component of large genomes in *Pinus elliottii* var. *elliottii* and other gymnosperms. *Proceedings of the National Academy of Sciences of the USA* **93**, 2708–2713.
Kapitonov, V. V. & Jurka, J. (1999). Molecular paleontology of transposable elements from *Arabidopsis thaliana*. *Genetica* **107**, 27–37.
Katsiotis, A., Schmidt, T. & Heslop-Harrison, J. S. (1996). Chromosomal and genomic organization of Ty1-*copia*-like retrotransposon sequences in the genus *Avena*. *Genome* **39**, 410–417.
Keightley, P. D. (1994). The distribution of mutation effects on viability in *Drosophila melanogaster*. *Genetics* **138**, 1315–1322.
Kelly, J. K. (1999*a*). Response to selection in partially self-fertilizing populations. I. Selection on a single trait. *Evolution* **53**, 336–349.
Kelly, J. K. (1999*b*). Response to selection in partially self-fertilizing populations. II. Selection on mulitple traits. *Evolution* **53**, 350–357.
Kidwell, M. G. & Lisch, D. R. (2000). Transposable elements and host genome evolution. *Trends in Ecology and Evolution* **15**, 95–99.
Kimura, M. & Maruyama, T. (1966). The mutational load with epistatic gene interactions in fitness. *Genetics* **54**, 1337–1351.
Kondrashov, A. S. (1984). Deleterious mutations as an evolutionary factor. I. The advantage of recombination. *Genetical Research* **44**, 199–217.
Kondrashov, A. S. (1985). Deleterious mutation as an evolutionary factor. II. Facultative apomixis and selfing. *Genetics* **111**, 635–653.

Konieczny, A., Voytas, D. F., Cummings, M. P. & Ausubel, F. M. (1991). A superfamily of *Arabidopsis thaliana* retrotransposons. *Genetics* **127**, 801–809.
Kumar, A. (1996). The adventures of the Ty1-*copia* group of retrotransposons in plants. *Trends in Genetics* **12**, 41–43.
Kumar, A. & Bennetzen, J. L. (1999). Plant retrotransposons. *Annual Reviews of Genetics* **33**, 479–532.
Kumar, A., Pearce, S. R., McLean, K., Harrison, G., Heslop-Harrison, J. S., Waugh, R. & Flavell, A. J. (1997). The Ty1-*copia* group of retrotransposons in plants: genomic organisation, evolution, and use as molecular markers. *Genetica* **100**, 205–217.
Lande, R. (1976). Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30**, 314–334.
Lande, R., Schemske, D. W. & Schultz, S. T. (1994). High inbreeding depression, selective interference among loci, and the threshold selfing rate for purging recessive lethal mutations. *Evolution* **48**, 965–978.
Langley, C. H., Brookfield, J. F. Y. & Kaplan, N. (1983). Transposable elements in Mendelian populations. I. A theory. *Genetics* **104**, 457–471.
Langley, C. H., Montgomery, E., Hudson, R., Kaplan, N. & Charlesworth, B. (1988). On the role of unequal exchange in the containment of transposable element copy number. *Genetical Research* **52**, 223–235.
Laten, H. M. & Morris, R. O. (1993). SIRE-1, a long interspersed repetitive DNA element from soybean with weak sequence similarity to retrotransposons: initial characterization and partial sequence. *Gene* **134**, 153–159.
Lim, J. K. & Simmons, M. J. (1994). Gross chromosome rearrangements mediated by transposable elements in *Drosophila melanogaster*. *Bioessays* **16**, 269–275.
Lin, X., Kaul, S., Rounsley, S., Shea, T., Benito, M., Town, C., Fujii, C., Mason, T., Bowman, C., Barnstead, M., Feldblyum, T., Buell, C., Ketchum, K., Lee, J., Ronning, C., Koo, H., Moffat, K., Cronin, L. & Shen, M., *et al*.. (1999). Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature* **402**, 761–768.
Lonning, W.-E. & Saedler, H. (1997). Plant transposons: contributors to evolution? *Gene* **205** 245–253.
Lyman, R. F., Lawrence, F., Nuzhdin, S. V. & Mackay, T. F. C. (1996). Effects of single *P*-element insertions on bristle number and viability in *Drosophila melanogaster*. *Genetics* **143**, 277–292.
Mackay, T. F., Lyman, R. F. & Jackson, M. S. (1992). Effects of *P* element insertions on quantitative traits in *Drosophila melanogaster*. *Genetics* **130**, 315–332.
Mao, L., Wood, T. C., Yu, Y., Budiman, M. A., Tomkins, J., Woo, S., Sasinowski, M., Presting, G., Frisch, D., Goff, S., Dean, R. A. & Wing, R. A. (2000). Rice transposable elements: a survey of 73000 sequence-tagged-connectors. *Genome Research* **10**, 982–990.
Matzke, M. A. & Matzke, A. J. M. (1998). Polyploidy and transposons. *Trends in Ecology and Evolution* **13**, 241.
McClintock, B. (1984). The significance of responses of the genome to challenge. *Science* **226**, 792–801.
McDonald, J. (1995). Transposable elements: possible catalysts of organismic evolution. *Trends in Ecology and Evolution* **10**, 123–126.
McFadden, J. & Knowles, G. (1997). Escape from evolutionary stasis by transposon-mediated deleterious mutations. *Journal of Theoretical Biology* **186**, 441–447.
Mhiri, C., Morel, J. B., Vernhettes, S., Casacuberta, J. M., Lucas, H. & Grandbastien, M. A. (1997). The promoter of the tobacco *Tnt1* retrotransposon is induced by wounding and by abiotic stress. *Plant Molecular Biology* **33**, 257–266.
Miller, J. T., Dong, F., Jackson, S. A., Song, J. & Jiang, J. (1998). Retrotransposon-related DNA sequences in the centromeres of grass chromosomes. *Genetics* **150**, 1615–1623.
Montgomery, E., Charlesworth, B. & Langley, C. H. (1987). A test for the role of natural selection in the stabilization of transposable element copy number in a population of *Drosophila melanogaster*. *Genetical Research* **49**, 31–41.
Montgomery, E. A., Huang, S. M., Langley, C. H. & Judd, B. H. (1991). Chromosome rearrangement by ectopic recombination in *Drosophila melanogaster*: genome structure and evolution. *Genetics* **129**, 1085–1098.
Moore, G., Lucas, H., Batty, N. & Flavell, R. (1991). A family of retrotransposons and associated genomic variation in wheat. *Genomics* **10**, 461–468.
Nakajima, R., Noma, K., Ohtsubo, H. & Ohtsubo, E. (1996). Identification and characterization of two tandem repeat sequences (TrsB and TrsC) and a retrotransposon (RIRE1) as genome-general sequences in rice. *Genes and Genetic Systems* **71**, 373–382.
Nevers, P. & Saedler, H. (1977). Transposable genetic elements as agents of gene instability and chromosomal rearrangements. *Nature* **268**, 109–115.
Noma, K., Nakajima, R., Ohtsubo, H. & Ohtsubo, E. (1997). RIRE1, a retrotransposon from wild rice *Oryza australiensis*. *Genes and Genetic Systems* **72**, 131–140.
Nuzhdin, S. V. (1999). Sure facts, speculations, and open questions about the evolution of transposable element copy number. *Genetica* **107**, 129–137.
Nuzhdin, S. V. & Mackay, T. F. C. (1994). Direct determination of retrotransposon transposition rates in *Drosophila melanogaster*. *Genetical Research* **63**, 139–144.
Nuzhdin, S. V. & Mackay, T. F. (1995). The genomic rate of transposable element movement in *Drosophila melanogaster*. *Molecular Biology and Evolution* **12**, 180–191.
Nuzhdin, S. V., Pasyukova, E. G. & Mackay, T. F. (1997). Accumulation of transposable elements in laboratory lines of *Drosophila melanogaster*. *Genetica* **100**, 167–175.
Pearce, S. R., Harrison, G., Li, D., Heslop-Harrison, J. S., Kumar, A. & Flavell, A. J. (1996*a*). The Ty1-*copia* group retrotransposons in *Vicia* species: copy number, sequence heterogeneity and chromosomal localisation. *Molecular and General Genetics* **250**, 305–315.
Pearce, S. R., Pich, U., Harrison, G., Flavell, A. J., Heslop Harrison, J. S. P., Schubert, I. & Kumar, A. (1996*b*). The Ty1-*copia* group retrotransposons of *Allium cepa* are distributed throughout the chromosomes but are enriched in the terminal heterochromatin. *Chromosome Research* **4**, 357–364.
Pearce, S. R., Harrison, G., Heslop-Harrison, P. J., Flavell, A. J. & Kumar, A. (1997). Characterization and genomic organization of Ty1-*copia* group retrotransposons in rye (*Secale cereale*). *Genome* **40**, 617–625.
Pouteau, S., Huttner, E., Grandbastien, M. A. & Caboche, M. (1991). Specific expression of the tobacco Tnt1 retrotransposon in protoplasts. *EMBO Journal* **10**, 1911–1918.
SanMiguel, P., Tikhonov, A., Jin, Y.-K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., Springer, P. S., Edwards, K. J., Lee, M., Avramova, Z. & Bennetzen, J. L. (1996). Nested retrotransposons in the intergenic regions of the maize genome. *Science* **274**, 765–768.
Sant, V. J., Sainani, M. N., Sami-Subbu, R., Ranjekar, P. K. & Gupta, V. S. (2000). Ty1-*copia* retrotransposon-like elements in chickpea genome: their identification, distribution and use for diversity analysis. *Gene* **257**, 157–166.
Schmidt, T., Kubis, S. & Heslop-Harrison, J. S. (1995).

Analysis and chromosomal localization of retrotransposons in sugar beet (*Beta vulgaris* L.): LINEs and Ty1-*copia*-like elements as major components of the genome. *Chromosome Research* **3**, 335–345.

Shalev, G. & Levy, A. A. (1997). The maize transposable element *Ac* induces recombination between the donor site and an homologous ectopic sequence. *Genetics* **146**, 1143–1151.

Suoniemi, A., Anamthawat-Jonsson, K., Arna, T. & Schulman, A. H. (1996). Retrotransposon BARE-1 is a major, dispersed component of the barley (*Hordeum vulgare* L.) genome. *Plant Molecular Biology* **30**, 1321–1329.

Suoniemi, A., Tanskanen, J. & Schulman, A. H. (1998). *Gypsy*-like retrotransposons are widespread in the plant kingdom. *Plant Journal* **13**, 699–705.

The Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796–815.

Tsitrone, A., Charles, S. & Biémon, C. (1999). Dynamics of transposable elements under the selection model. *Genetical Research* **74**, 159–164.

Turcich, M. P., BokhariRiza, A., Hamilton, D. A., He, C. P., Messier, W., Stewart, C. B. & Mascarenhas, J. P. (1996). Prem-2, a *copia*-type retroelement in maize is expressed preferentially in early microspores. *Sexual Plant Reproduction* **9**, 65–74.

VanderWiel, P. L., Voytas, D. F. & Wendel, J. F. (1993). *Copia*-like retrotransposable element evolution in diploid and polyploid cotton (*Gossypium* L.). *Journal of Molecular Evolution* **36**, 429–447.

Verries, C., Bes, C., This, P. & Tesniere, C. (2000). Cloning and characterization of Vine-1, a LTR-retrotransposon-like element in *Vitis vinifera* L., and other *Vitis* species. *Genome* **43**, 366–376.

Vicient, C. M., Kalendar, R., Anamthawat-Jonsson, K. & Schulman, A. H. (1999). Structure, functionality, and evolution of the BARE-1 retrotransposon of barley. *Genetica* **107**, 53–63.

Voytas, D. F. (1996). Retroelements in genome organization. *Science* **274**, 737–738.

Voytas, D. F. & Ausubel, F. M. (1988). A *copia*-like transposable element family in *Arabidopsis thaliana*. *Nature* **336**, 242–244.

Voytas, D. F., Konieczny, A., Cummings, M. P. & Ausubel, F. M. (1990). The structure, distribution and evolution of the *Ta1* retrotransposable element family of *Arabidopsis thaliana*. *Genetics* **126**, 713–721.

Voytas, D. F., Cummings, M. P., Konieczny, A., Ausubel, F. M. & Rodermel, S. R. (1992). *Copia*-like retrotransposons are ubiquitous among plants. *Proceedings of the National Academy of Sciences of the USA* **89**, 7124–7128.

Wang, S., Zhang, Q., Maughan, P. J. & Saghai Maroof, M. A. (1997). *Copia*-like retrotransposons in rice: sequence heterogeneity, species distribution and chromosomal locations. *Plant Molecular Biology* **33**, 1051–1058.

Weil, C. F. & Wessler, S. R. (1990). The effects of plant transposable element insertion on transcription initiation and RNA processing. *Annual Review of Plant Physiology and Plant Molecular Biology* **41**, 527–552.

Wessler, S. R. (1996). Plant retrotransposons: turned on by stress. *Current Biology* **6**, 959–961.

Wessler, S. R., Bureau, T. E. & White, S. E. (1995). LTR-retrotransposons and MITES: important players in the evolution of plant genomes. *Current Opinion in Genetics and Development* **5**, 814–821.

White, S. E., Habera, L. F. & Wessler, S. R. (1994). Retrotransposons in the flanking regions of normal plant genes: a role for *copia*-like elements in the evolution of gene structure and expression. *Proceedings of the National Academy of Sciences of the USA* **91**, 11792–11796.

Wilke, C. M., Maimer, E. & Adams, J. (1992). The population biology and evolutionary significance of *Ty* elements in *Saccharomyces cerevisiae*. *Genetica* **86**, 155–173.

Wright, D., Ke, N., Smalle, J., Hauge, B., Goodman, H. & Voytas, D. (1996). Multiple non-LTR retrotransposons in the genome of *Arabidopsis thaliana*. *Genetics* **142**, 569–578.

Wright, S. I. & Schoen, D. J. (1999). Transposon dynamics and the breeding system. *Genetica* **107**, 139–148.

Yanez, M., Verdugo, I., Rodriguez, M., Prat, S. & Ruiz-Lara, S. (1998). Highly heterogeneous families of Ty1/*copia* retrotransposons in the *Lycopersicon chilense* genome. *Gene* **222**, 223–228.

Young, R., Francis, D., St Clair, D. & Taylor, B. (1994). A dispersed family of repetitive DNA sequences exhibits characteristics of a transposable element in the genus *Lycopersicon*. *Genetics* **137**, 581–588.