Ⓔ

The Evolution of Recombination under Domestication:
A Test of Two Hypotheses

Jeffrey Ross-Ibarra[*]

Department of Genetics, Life Sciences Building, University of Georgia, Athens, Georgia 30602

ABSTRACT: The successful domestication of wild plants has been one of the most important human accomplishments of the last 10,000 yr. Though our empirical knowledge of the genetic mechanisms of plant domestication is still relatively limited, there exists a large body of theory that offers a host of hypotheses on the genetics of domestication. Two of these that have not been addressed concern the role of recombination in the process of domestication. The first predicts an increase in recombination rate through domestication, while the second argues that recombination rate should serve as a preadaptation to domestication. This study makes use of data on chiasma frequencies available from almost a century of plant cytogenetical literature to test these two hypotheses. The results support the hypothesis that domestication selects for an increase in recombination, and in rejecting the preadaptation hypothesis, they suggest directions for future research into the possibility of preadaptation to domestication.

Keywords: domestication, recombination, preadaptation, chiasma frequency.

The successful domestication of wild plants has been one of the most important human accomplishments of the last 10,000 yr. Changes wrought by domestication enabled human populations to harness and control a food supply tremendously greater than was previously possible. The magnitude of these changes and the rapidity with which they were effected are convincing evidence of the strong directional selection to which these plants were subjected. Though clearly each species is a product of its own unique history, several patterns of morphological and genetic

* E-mail: jri@uga.edu.

change can nonetheless be discerned among many domesticated plant species (Harlan 1992; Paterson 2002). Morphological and physiological changes including gigantism, loss of natural dispersal and defense mechanisms, and loss of seed dormancy have been known for decades and are well documented for a wide variety of crop plants (Smartt and Simmonds 1995). Genetic evidence of strong selective pressure (e.g., Wang et al. 1999), loss of genetic diversity (Doebley 1989), and polyploidy (Hilu 1993) have been shown for some of the more important crop plants, but our empirical knowledge of the genetics behind the domestication process is still rather limited.

Though the empirical literature on domestication is less developed than could be desired, there exists a large body of theory on the genetics of plant domestication. Both quantitative and population genetics provide a host of hypotheses regarding domestication. Two of these that have not been addressed concern the role of recombination in the process of domestication.

The first, initially proposed by Rees and Dale (1974) and later echoed by Burt and Bell (1987) and Otto and Barton (2001), predicts an increase in recombination rate through domestication. Both theory and simulations show that selection generally favors an increased recombination rate during periods of rapid evolutionary change (Otto and Barton 1997). High recombination is of most value when selection is strong and genetic variability is limited by negative linkage disequilibrium (Feldman et al. 1997). In domesticated plants this disequilibrium might have been generated by population bottlenecks and genetic drift (Felsenstein and Yokoyama 1976; Otto and Barton 1997, 2001) or by negative epistasis among beneficial alleles (Charlesworth 1993; Barton 1995).

Gornall (1983) elaborated the second hypothesis, arguing that high recombination rate should serve as a preadaptation to domestication: because high recombination rate increases response to strong selection, he argued, plants without this advantage would be less likely to be successfully domesticated.

The two hypotheses are not mutually exclusive but, while based on very similar theoretical underpinnings,

make distinct and readily testable predictions about patterns of recombination rate in domesticated plants and their wild relatives. Rees and Dale's hypothesis predicts a higher recombination rate in domesticated plants relative to their wild progenitors, while Gornall's hypothesis predicts a higher recombination rate in the wild progenitors of domesticated plants relative to their other wild congeners.

This study makes use of data on chiasma frequencies available from almost a century of plant cytogenetical literature to test these two hypotheses. The results not only support the predictions of Rees and Dale and others but also, in rejecting Gornall's hypothesis, suggest directions for future research into the possibility of preadaptation to domestication.

Methods

Recombination data were collected from published studies citing chiasmata frequencies for species of vascular plants. Though it has been suggested that chiasma frequency is not always an accurate measure of recombination in plants (Nilsson et al. 1993), recent studies have shown the direct one-to-one relation between chiasmata and recombination events (King et al. 2002; Knox and Ellis 2002). Data on other characteristics were obtained from electronic databases, floras, and a variety of other sources. A text file of the recombination data used is available as table B1 in the online edition of the *American Naturalist*, and a complete list of the sources consulted is available on request to the author.

Five parameters were recorded for the recombination data: haploid chromosome number (N), total number of chiasmata per nucleus (XMA), excess chiasmata per nucleus ($XS = XMA - N$), chiasmata per bivalent ($II = XMA/N$), and recombination index ($RI = XMA + N$). Missing data were calculated using the above relationships or, in the case of haploid number, taken from the Index to Plant Chromosome Numbers (IPCN) database (2002) or from Darlington and Wylie (1956). Data from each study were entered into the database as a species average weighted by the number of plants or cells of each species examined in the study. Species data from multiple publications were averaged to give a single value in the final data set. If entries for the same species differed in one of the other characteristics recorded, however, those entries were not averaged but instead kept separate in the data set. Data were recorded separately for male and female meioses; only data from male meiosis were included in this study. Whenever possible, chiasma frequencies at metaphase were used. Likewise, data from plants with supernumerary chromosomes were excluded whenever possible. Nonetheless, sources that included only data from other phases of meiosis or from plants with supernumerary chromosomes were included in the database.

Each species was classified for five additional characteristics: ploidy level, life form (annual or perennial), mating system (selfing, mixed, or outcrossing), weediness (weed or not), and domestication status (wild, cultivated, or domesticated). Data on ploidy level, life form, and mating system were occasionally available from the original source but were otherwise gleaned from a wide range of sources. Plants were considered weeds if named as such by three or more sources as cited in the Global Compendium of Weeds (2002). The domestication status provided by the original source was used whenever available. If not available, various sources were consulted to determine the appropriate status; the majority of the determinations, however, were made using Smartt and Simmonds (1995). Plants grown horticulturally and plants for which no clear evidence of domestication was available (mostly forage grasses and relatives of domesticated taxa) were considered cultivated but not domesticated for purposes of this study.

Statistical analyses took the form of an across-species ANOVA of the entire data set followed by pairwise comparison of specific groups. The raw data for haploid number and chiasma frequency were significantly nonnormally distributed and were transformed using a Box-Cox transformation for use in the ANOVA and pairwise tests.

To test the importance of domestication in determining recombination rate, a stepwise regression analysis was performed on the entire data set. Life form, mating system, weediness, domestication, and all two-way interactions were included as variables in the initial model, but all three- and four-way interactions were ignored. More direct tests of the effect of domestication were carried out by making pairwise comparisons of domesticated species to their wild progenitors and of the progenitors to their nearest congeners. These were compared using both a standard paired *t*-test as well as a Wilcoxon paired-samples test of the untransformed data. To test for possible correlations between these pairwise differences and the other species characteristics, the sets of (transformed) pairwise differences were then analyzed by two stepwise regressions, again ignoring all three- and four-way interaction terms. The first regression analyzed the effect of the other species characteristics on the size of the difference in recombination rate between paired taxa (e.g., whether selfing taxa have smaller differences than outcrossing taxa). The second looked for correlations between a change in these species characteristics as a result of domestication and the difference in recombination rate (e.g., whether a change from selfing to outcrossing is correlated with larger differences in recombination than no change in mating system). Statistical calculations were carried out using Statistica 5.5 (Statsoft 2000).

Table 1: Regression analysis of entire data set for chiasmata per bivalent

	df	SS	MS	*F*	*P*
Intercept	1	28.347	28.347	322.955	.000
Domestication	2	.552	.276	3.14	.045
Mating	2	.637	.319	3.63	.028
Weediness	1	.028	.028	.318	.573
Life form	1	.056	.056	.634	.427
Weediness × life form	1	.434	.434	4.941	.027
Error	191	16.765	.088		

Note: The following factors showing no data were removed from the model because their corresponding *P* values were >.1: domestication × weediness, domestication × life form, domestication × mating, weediness × mating, life form × mating.

Results

Data were collected for 601 species of vascular plants from 124 genera and 37 families. After elimination of incomplete cases, the across-species analyses of chiasma frequency were limited to a sample size of 196, including 46 domesticated species. Species used for pairwise comparisons are listed in table A1 in the online edition of the *American Naturalist.*

In order to separate the effect of chromosome number in determining chiasma frequency, it was desirable to use a measure of chiasma frequency independent of haploid number and ploidy level. In spite of claims that excess chiasmata is independent of chromosome number (Burt and Bell 1987; Koella 1993), all measures of chiasma frequency in this data set are significantly correlated with haploid number (Pearson product-moment correlation of

>0.5 for all measures besides chiasmata per bivalent, significant at the $P < .05$ level), and all but chiasmata per bivalent are significantly influenced by ploidy level ($P < .0001$ in all cases). Chiasmata per bivalent is least correlated with haploid number (Pearson product-moment correlation of -0.20) and is independent of ploidy level ($P > .27$). Though only data using chiasmata per bivalent are reported here, analyses performed using the other measures do not differ qualitatively from those presented.

Results from an initial across-species regression analysis are shown in table 1. Though domestication is not the sole determinant of recombination rate in the final model, it is clearly significant. A post hoc analysis reveals that although cultivated and wild plants are not discernibly different from each other, domesticated plants have a significantly higher recombination rate than either of the former (Scheffé test, $P < .03$ for wild and $P < .01$ for cultivated), providing support for the hypothesis that domestication increases recombination rate. It is worth noting that none of the other characteristics (mating system, life form, etc.) interact significantly with domestication, suggesting that the role these have played in the effect of domestication on recombination rate is relatively insignificant. Including the progenitors of crop plants in the regression model as a category distinct from other wild plants adds no meaningful information; the relative sample size of the category is small and the standard error such that it cannot be distinguished from either other wild or domesticated taxa.

By performing an across-taxa regression for genera and families, the data make possible a test for preadaptation
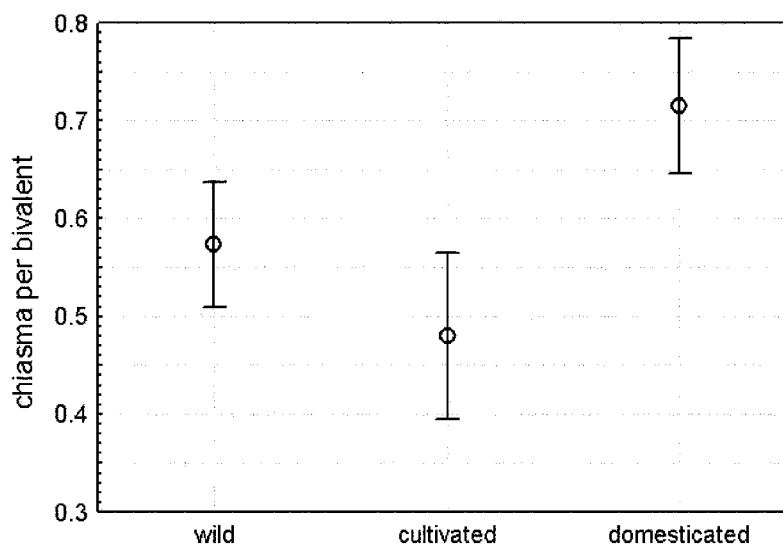

Figure 1: Observed means and 95% confidence intervals of the three categories of domestication; transformed data shown
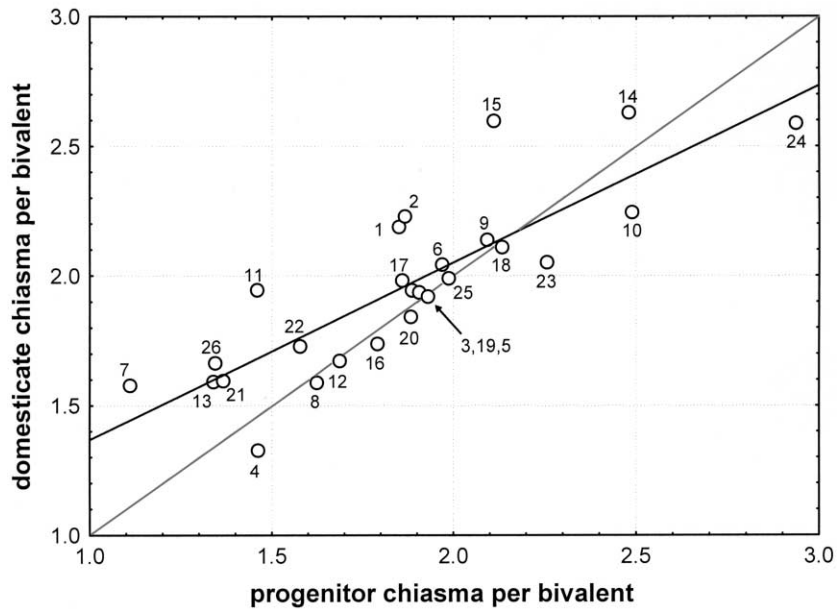
Figure 2: Pairwise comparison of domesticated taxa to their nearest congener. In each case the chiasmata per bivalent of the domesticate is plotted against the chiasmata per bivalent of the progenitor. Numbers refer to species in table A1 in the online edition of the *American Naturalist*.
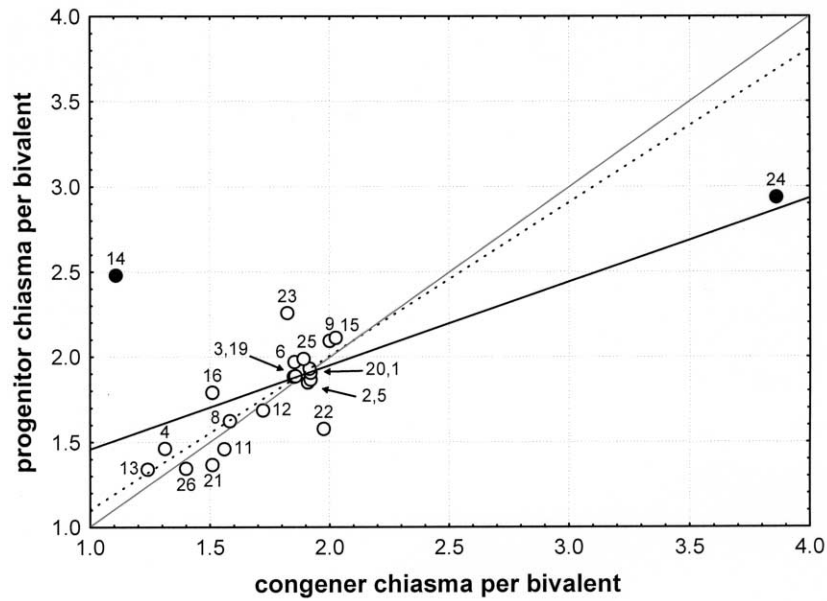

Figure 3: Pairwise comparison of progenitor taxa to their nearest congener. In each case the chiasmata per bivalent of the progenitor is plotted against the chiasmata per bivalent of the congener. *Solid line*, regression when all data points are included. *Dotted line*, regression when the two outlier taxa (*filled circles*) are removed from the analysis. Numbers refer to species in table A1 in the online edition of the *American Naturalist*.

Table 2: Number of species pairs, taxa excluded, average difference in chiasma per bivalent and its SE, and *P* values associated with domesticate-progenitor (D-P) and progenitor-congener (P-C) pairwise comparisons

	D-P	D-P	D-P	P-C	P-C
Species pairs	26	22	20	22	20
Taxa excluded	…	7,10,17,18	7,10,14 17,18,24	7,10,17,18	7,10,14 17,18,24
Average difference (SE)	.095 (.044)	.097 (.046)	.117 (.045)	.048 (.083)	.031 (.037)
Paired *t*	.02	.023	.009	.283	.208
Wilcoxon	.028	.029	.017	.19	.161

Note: A paired *t*-test was performed on the transformed data, and a Wilcoxon paired-samples test was calculated using nontransformed data. The numbers of taxa excluded in each comparison refer to table A1 in the online edition of the *American Naturalist*, and the rationalization for each comparison is explained in the text. The mean SD of replicate measurements of chiasma per bivalent within a species was .140.

at these taxonomic levels. Were preadaptation acting at these higher levels, genera or families that gave rise to domesticated taxa would be expected to have higher recombination rates than genera or families that did not. Such an analysis reveals no significant effect ($P > .7$ for both families and genera; results do not differ if data from domesticated species are included in the analysis), suggesting that preadaptation is not an important factor at the genus or family level.

A more direct test of the role of recombination is a pairwise test of the actual taxa involved. These are shown graphically in figures 2 and 3, and *P* values for these tests are reported in table 2. Though the overall data conform reasonably well to normality after transformation, the set of transformed pairwise comparisons (both the raw data and the set of differences themselves) is still markedly nonnormal. While the *t*-test is generally robust to nonnormality, a nonparametric test is perhaps more appropriate for these analyses; results for both a standard paired *t*-test and the Wilcoxon paired-samples test are reported in table 2.

The comparison of domesticates to their wild progenitors is shown in figure 2. Domesticated taxa in general show a significant increase in recombination when compared to their wild relatives (table 2). Hexaploid domesticated oats (*Avena sativa*) differ in ploidy from their most direct wild progenitor, the tetraploid *Avena insularis*. Given that chiasmata per bivalent is independent of ploidy level, this difference should not affect the results; however, even if *Avena* is removed from the list, the results do not change meaningfully.

Of the 26 species pairs from table A1 compared above, recombination data for both a progenitor and a congener were available in only 22 cases. For three of the progenitor taxa (*Cucumis*, *Pennisetum*, and *Triticum*), the closest available congener was of different ploidy level; removal of these taxa once again does not change the results. Two cases (*Oryza* and *Vicia*) were found to be statistical outliers

(Grubb's test for outliers, $P < .05$). In both cases, the congeners used, though the closest relative of the progenitor, are dramatically different from other species in their genera and probably should be excluded as misrepresentative. Nonetheless, results from both comparisons (the original 22 and the 20 nonoutliers) are shown in table 2 and are plotted on figure 3. The comparison does not detect a significant difference between progenitors and their congeners, thus failing to support Gornall's hypothesis of preadaptation. Though the number of taxa analyzed is small, it is unlikely that any large difference was missed by this test. Using the same reduced set of 22 or 20 taxa to make the paired domesticate-progenitor comparison still produces a statistically significant result (table 2), indicating that even if recombination rate is of preadaptive value, its effect is less than that of domestication in increasing recombination rate via selection. Moreover, a post hoc power analysis (Hintze 2001) based on the 20 progenitor-congener pairs does not reveal any real lack of power (80% power to detect a difference of 0.061 in the transformed data, a difference less than that found in any of the comparisons of domesticates and progenitors).

Finally, data on mating system (outcrossing, mixed, inbreeding), weediness, and life form (annual vs. perennial) were used to determine whether these characters, or changes in these characters, are correlated with differences found in the pairwise tests. None of the ANOVAs of pairwise differences in recombination rate with regard to these species characteristics or changes in these characteristics was significant ($P > .10$ in all cases) for either of the comparisons (domesticate-progenitor or progenitor-congener). The increase in recombination rate due to domestication does not seem to be affected by any of the other traits studied or changes in these traits, nor does the lack of preadaptation seem to be explained by these other characters.

Discussion

The results from both an across-species analysis and a paired comparison of domesticates to their progenitor taxa strongly suggest that the domestication process generally increases the recombination rate of a species. Domesticated taxa show a higher overall recombination rate than nondomesticated taxa (fig. 1), and pairwise comparison to their progenitors reaffirms the result. Though chiasmata per bivalent is undoubtedly the most appropriate measure of recombination for this analysis, it is comforting to note that all measures tested showed similar results.

In spite of the data available, the literature includes few observations that domestication might affect chiasma frequencies. Though several studies have published data on both a domesticate and its progenitor, very few authors have noted the difference in recombination rate (but see Koul et al. 1989), and none have offered a plausible explanation. In addition to these few empirical papers, a study by Burt and Bell (1987) on recombination rate and life span in mammals has occasionally been cited as evidence of the effect of domestication on recombination rate (Koul et al. 1989; Otto and Barton 2001). While their data do show a statistically significant increase in excess chiasma among several domesticated mammal species, their study does not include any of the progenitors of these species, and the effect disappears if the analysis is done using measures of recombination shown to be less strongly correlated with chromosome number (data not shown). With the exception of the limited work mentioned above, then, this study is the first to conclusively show the general effect of domestication on recombination rate.

In contrast to the lack of empirical work, there is a large body of theory that would predict a correlation between domestication and a change in recombination rate. Much effort has been devoted to determining the potential sources of the negative linkage disequilibria and the conditions under which recombination is favored. It is widely agreed, however, that strong directional selection (Feldman et al. 1997), especially at multiple loci (Otto and Barton 1997) or in concert with genetic drift (Felsenstein and Yokoyama 1976; Otto and Barton 2001), can generate negative disequilibria sufficient to select for increased recombination. And while it has been suggested that introgression from wild relatives could potentially select for a decreased recombination rate in domesticates (Lenormand and Otto 2000), the majority of the conditions provided by the process of domestication (new environment, strong directional selection at multiple loci, and, at least in some cases, small population sizes) concur with those thought to select for recombination.

Theoretical analysis offers many scenarios for the evolution of recombination; it also cautions of the negative consequences of recombination rate. Not only does recombination break down negative disequilibria among beneficial alleles, but it also disrupts positive disequilibria among them as well, potentially splitting up adaptive gene complexes (Barton 1995). This "recombination load" is thought to limit selection for increased recombination and effectively set a threshold above which higher recombination is selected against. Selection against excessive recombination would predict a regression slope of <1 between the recombination rate of a domesticate and that of its progenitor. Such a relation is clearly seen in figure 2 ($P < .05$). Interestingly, the regression line crosses the 1 : 1 line of equality at just above 2 chiasmata/bivalent, a number suggested by Kondrashov (1988) as a potential limit beyond which the effects of recombination become detrimental. Perhaps most significantly, a slope of <1 is not in keeping with an alternative explanation for the association between domestication and high recombination rate: the idea that increased homozygosity, and not selection, is the primary cause of increased recombination.

While the results obtained strongly support the hypothesis that domestication selects for increased recombination rate, they offer no support for the hypothesis that recombination rate is an important preadaptation to domestication. Pairwise comparisons reveal that not only do the wild progenitors of crop plants generally possess lower recombination rates than their domesticated descendants but also that they are not discernibly different from their wild congeners. Furthermore, if recombination functions as a preadaptation to domestication, constraints due to recombination load should be evident as recombination increases in the congeners of domesticated species. As recombination increases in congeneric taxa, selection in favor of progenitors with higher recombination should decrease, resulting in a slope of <1 as seen in figure 2. Though this relation is in fact seen in figure 3, the removal of the two outlier taxa reveals that it is probably artifactual; the remaining tightly-clustered comparisons do not reveal any such relation (the regression line is not different from 1). Likewise, across-taxa comparisons at the genus and family levels similarly find no effect of preadaptation, and there are no correlations with mating system, life form, or weediness that explain the lack of an effect.

Though we are beginning to form an idea of the effect that domestication has had on the genetics of plant species, we are still far from understanding the role genetics has played in determining which species were successfully domesticated. It seems implausible that ecological factors alone (mating system, life form) could explain the impressive discrepancy between the 250,000 species of flowering plants and the few hundred species of domesticates (Hawkes 1983; Diamond 2002). Yet the very extent to which humans have modified or utilized the botanical di-

versity of their environments almost requires a mechanistic explanation for this discrepancy, and it seems quite plausible that genetic factors could have played a central role in the success or failure of the domestication process. While the genetic bases of many important domestication traits have been identified (Paterson 2002), we will not know how these genes have influenced the success of domestication until they have been studied in many wild species as well. Other hypothesized genetic preadaptations, such as polyploidy (Hilu 1993), have been shown to be unimportant in determining the successful domestication of plant species, and the present analysis suggests that recombination rate is likewise of little importance. Clearly, much work is still needed before we can approach a more complete understanding of the genetic mechanisms involved in the domestication process.

Acknowledgments

Literature Cited

Barton, N. H. 1995. A general model for the evolution of recombination. Genetical Research 65:123–144.

Burt, A., and G. Bell. 1987. Mammalian chiasma frequencies as a test of two theories of recombination. Nature 326:803–805.

Charlesworth, B. 1993. Directional selection and the evolution of sex and recombination. Genetical Research 61:205–224.

Darlington, C. D., and A. P. Wylie. 1956. Chromosome atlas of flowering plants. Macmillan, New York.

Diamond, J. 2002. Evolution, consequences, and future of plant and animal domestication. Nature 418:700–707.

Doebley, J. F. 1989. Isozymic evidence and the evolution of crop plants. Pages 165–191 *in* D. E. Soltis and P. S. Soltis, eds. Isozymes in plant biology. Dioscorides, Portland, Oreg.

Feldman, M. W., S. P. Otto, and F. B. Christiansen. 1997. Population genetic perspectives on the evolution of recombination. Annual Review of Genetics 30:261–295.

Felsenstein, J., and S. Yokoyama. 1976. The evolutionary advantage of recombination. II. Individual selection for recombination. Genetics 83:845–859.

A global compendium of weeds. 2002. http://www.hear.org/gcw/index.html.

Gornall, R. J. 1983. Recombination systems and plant domestication. Biological Journal of the Linnean Society 20:375–383.

Harlan, J. R. 1992. Crops and man. American Society of Agronomy/Crop Science Society of America, Madison, Wis.

Hawkes, J. G. 1983. The diversity of crop plants. Harvard University Press, Cambridge, Mass.

Hilu, K. W. 1993. Polyploidy and the evolution of domesticated plants. American Journal of Botany 80:1494–1499.

Hintze, J. 2001. NCSS and PASS. Number Cruncher Statistical Systems, Kaysville, Utah.

Index to plant chromosome numbers. 2002. http://mobot.mobot.org/W3T/Search/ipcn.html.

King, J., L. A. Roberts, M. J. Kearsey, H. M. Thomas, R. N. Jones, L. Huang, I. P. Armstead, W.G. Morgan, and I.P King. 2002. A demonstration of a 1 : 1 correspondence between chiasma frequency and recombination using a *Lolium perenne/Festuca pratensis* substitution. Genetics 161:307–314.

Knox, M. R., and T. H. N. Ellis. 2002. Excess heterozygosity contributes to genetic map expansion in pea recombinant inbred populations. Genetics 162:861–873.

Koella, J. C. 1993. Ecological correlates of chiasma frequency and recombination index of plants. Biological Journal of the Linnean Society 48:227–238.

Kondrashov, A. S. 1988. Deleterious mutations and the evolution of sexual reproduction. Nature 336:435–440.

Koul, P., A. K. Koul, and I. A. Hamal. 1989. Reproductive biology of wild and cultivated carrot *Daucus carota* L. New Phytologist 112:437–444.

Lenormand, T., and S. P. Otto. 2000. The evolution of recombination in a heterogeneous environment. Genetics 156:423–428.

Nilsson, N. O., T. Sall, and B. O. Bengtsson. 1993. Chiasma and recombination data in plants: are they compatible? Trends in Genetics 9:344–348.

Otto, S. P., and N. H. Barton. 1997. The evolution of recombination: removing the limits to natural selection. Genetics 147:879–906.

———. 2001. Selection for recombination in small populations. Evolution 55:1921–1931.

Paterson, A. 2002. What has QTL mapping taught us about plant domestication? New Phytologist 154:591–608.

Rees, H., and P. J. Dale. 1974. Chiasmata and variability in *Lolium* and *Festuca* populations. Chromosoma 47:335.

Smartt, J., and N. W. Simmonds. 1995. Evolution of crop plants. Longman, Harlow, Essex.

Stasoft. 2000. STATISTICA for Windows. Statsoft, Tulsa, Okla.

Wang, R.-L., A. Stec, J. Hey, L. Lukens, and J. Doebley. 1999. The limits of selection during maize domestication. Nature 398:236–239.

Associate Editor: John Willis