SHORT COMMUNICATION
Genome size and recombination in angiosperms: a second look

J. ROSS-IBARRA

Department of Genetics, University of Georgia, Athens, GA, USA

Abstract

Despite dramatic differences in genome size – and thus space for recombination to occur – previous workers found no correlation between recombination rate and genome size in flowering plants. Here I re-investigate these claims using phylogenetic comparative methods to test a large data set of recombination data in angiosperms. I show that genome size is significantly correlated with recombination rate across a wide sampling of species and that change in genome size explains a meaningful proportion (~20%) of variation in recombination rate. I show that the strength of this correlation is comparable with that of several characters previously linked to evolutionary change in recombination rate, but argue that consideration of processes of genome size change likely make the observed correlation a conservative estimate. And finally, although I find that recombination rate increases less than proportionally to change in genome size, several mechanistic and theoretical arguments suggest that this result is not unexpected.

Introduction

Recombination and genome size are highly labile characteristics of plant genomes. Variation in recombination rate is seen among genes within plant genomes (Tenaillon *et al.*, 2002) as well as among individuals within a population and populations within a species (Rees & Dale, 1974). Moreover, recombination rate has been shown to respond to direct or indirect selection over short periods of time (Harinarayana & Murty, 1971; Ross-Ibarra, 2004), and putative adaptive correlations have been documented with several life-history characteristics (Koella, 1993). Variation in genome size among angiosperms is even more impressive: flowering plant genome sizes span a range of more than three orders of magnitude (Bennett & Leitch, 2005), and dramatic variation in genome size is more than evident even among congeners (Rees & Durrant, 1986; Narayan & McIntyre, 1989; Jakob *et al.*, 2004).

Given the dramatic differences in genome size – and thus space in which recombination could occur – and the labile nature of recombination, a positive correlation between genome size and recombination rate per chro-

Correspondence: J. Ross-Ibarra, Department of Ecology and Evolution, 321 Steinhaus Hall, University of California, Irvine, CA 92697, USA.
Tel.: +1-949-678-9006; fax: 1-949-824-2181;
e-mail: rossibarra@gmail.com

mosome would seem natural. Abundant evidence for a strong within-genome correlation between chromosome size and recombination rate strengthens this hypothesis (Anderson *et al.*, 2001, 2003). It is thus somewhat surprising that all investigations of this relationship among species have found either no correlation (Cavalier-Smith, 1985; Rees & Durrant, 1986; Anderson *et al.*, 2001) or a negative one (Narayan & McIntyre, 1989). These findings, however, have been based on limited data sets from a few genera or widely unrelated taxa, and have failed to incorporate phylogenetic information into their analyses. Yet there is a strong phylogenetic component to the observed variation in genome size (Soltis *et al.*, 2003), and clear differences among genera and families for average levels of recombination (Rees & Dale, 1974).

Here I re-investigate findings that recombination rate is independent of genome size. Utilizing a data set of 279 Angiosperm species, I find that both standard linear regression and phylogenetic comparative methods reveal a significant positive correlation between genome size and recombination rate both across all species as well as within individual families and genera. I also show that the magnitude of the effect of genome size on recombination rate is similar to or greater than the effect of several life-history characteristics commonly linked to the evolution of recombination rates. Finally, however, I find that increases in recombination are not proportional

to increases in genome size, as might be predicted from purely spatial or mechanistic considerations, and that this relationship is robust to comparisons using only euchromatic or nonrepetitive fractions of the genome.

Materials and methods

I utilized a previously published data set on chiasmata counts and other plant characteristics (Ross-Ibarra, 2004), supplementing these data with additional life-history data from a range of sources. Haploid genome sizes (pg DNA) were extracted from the Kew Plant C value Database (Bennett & Leitch, 2005). Both genome size and recombination data were available for a subset of 279 species from 63 genera in 22 families; the complete data set is available upon request. To correct for the strong correlation between total chiasmata and chromosome number, I used chiasmata per bivalent as a measure of recombination; analyses using the residuals of a regression of total chiasmata on chromosome number gave qualitatively similar results, but are not presented. All analyses were performed using both total genome size and average bivalent size; only data on total genome size are reported for most analyses. Chiasmata and genome size data were natural log-transformed for all analyses.

I initially explored the relationship between genome size and recombination rate with standard linear regressions of chiasmata per bivalent on genome size for the entire data set as well as within each of the two largest families (Poaceae, Fabaceae) and four largest genera (*Senecio*, *Vicia*, *Allium* and *Lathyrus*). To evaluate the correlation between genome size and recombination rate in a phylogenetic context, I mapped haploid genome size and chiasmata per bivalent onto four different phylogenies (Fig. S1). I constructed the first phylogeny using the PHYLIP software package (Felsenstein, 1993) to generate a neighbour-joining tree from hand-aligned nucleotide sequence data of the chloroplast gene rbcL. I used a single GenBank rbcL sequence to represent each of 55 genera from the total data set (not all sequences used were from species represented in the data; a list of sequences is available upon request). Given the paucity of infrageneric data available for most taxa, I coded species' relationships within each genus as polytomies, assigned terminal branch lengths of 0.005, and then arbitrarily resolved polytomies into bifurcations by inserting branches of length 10^{-13}.

I constructed a second phylogeny using independent topological information from the literature and constraining all branches to a length of one. Deeper branches of the tree (family and above) came from Stevens (2005) and the Angiosperm Phylogeny Group (2003), whereas information for the shallower clades was taken from individual studies (Hsiao *et al.*, 1995; Olmstead *et al.*, 1999; Downie *et al.*, 2000; Hu *et al.*, 2000; Kajita *et al.*, 2001; Mason-Gamer *et al.*, 2002; Salamin *et al.*, 2002; Steele & Wojciechowski, 2003). I again assumed infra-
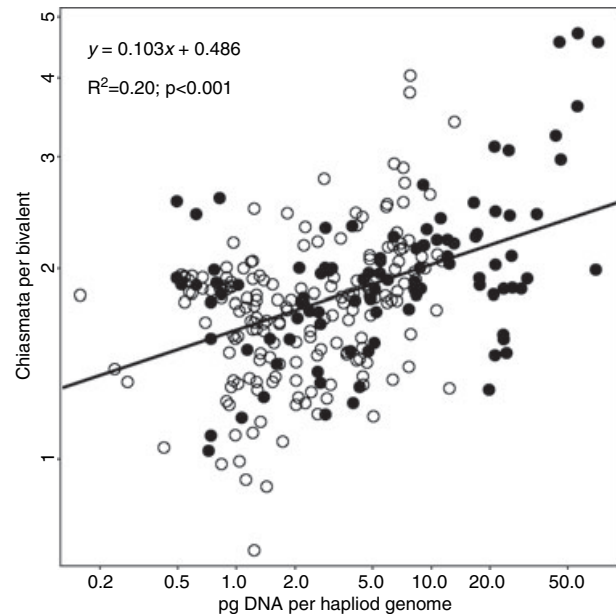

Fig. 1 Standard linear regression of chiasmata per bivalent on total genome size; solid circles represent monocotyledons and open circles dicotyledons. The regression equation displayed is for the combined log-transformed data.

generic polytomies which I arbitrarily divided into bifurcations by inserting branches of length 10^{-5}.

Finally, I constructed species-level trees for the two largest families in the data set (Poaceae and Fabaceae). Sequences from the 5.8S rDNA subunit and both internal transcribed spacers were available in GenBank for 29 species from each family. PHYLIP neighbour-joining trees were constructed from sequence alignments generated in CLUSTALW (Thompson *et al.*, 1994) and subsequently checked by hand.

Character evolution was analysed on each of the four phylogenies using the phylogenetic least squares model (PGLS; Martins & Hansen, 1997) implemented in the software COMPARE (Martins, 2004). In addition to correlations between characters, the model includes a parameter α which evaluates the strength of evolutionary constraint acting throughout the phylogeny. Under the assumption of α = 0 the PGLS becomes a standard model of independent contrasts (Felsenstein, 1985), and with an arbitrarily large α the model is equivalent to a non-phylogenetic linear regression. COMPARE provides regression results and a log-likelihood value for comparison of each of these alternative models.

To compare the effects of genome size on recombination to the effects of other factors linked to changes in recombination rate, I re-analysed the data using both of the two complete trees, including data on several other characters thought to be correlated with recombination: weediness (weedy, not weedy), domestication status (domesticated, cultivated and wild), mating system

(selfing, mixed and outcrossing) and perenniality (annual/biennial and perennial). Characters with more than two character states were recoded as a series of binary dummy variables (e.g. selfing vs. other, outcrossing vs. other). Data were not available for all species for these additional characters, and these analyses were performed on trees pruned to only include taxa with complete data (not shown). Comparisons using only weediness, domestication status and bivalent size, for which complete data were available, were also performed using the full tree.

Results and discussion

Consistent with a well-established pattern in angiosperms (Levin & Funderburg, 1979), no significant correlation was found between genome size and chromosome number ($r = -0.026$, $P = 0.66$) in these data. As expected, total genome size is thus tightly correlated ($r = 0.91$, $P < 0.001$) to average bivalent size, and comparisons of total genome size to recombination rate are qualitatively identical and quantitatively similar to comparisons made with average bivalent size. For most additional analyses, only results from comparisons to total genome size are therefore reported.

A significant positive correlation between recombination rate and genome size was observed in a standard regression analysis of all 279 species in the data set (Fig. 1). Although recombination rates were higher among monocotyledons ($t_{158} = 3.25$, $P < 0.01$), the slope of the regression of recombination rate and genome size was nearly identical in monocots and dicots (not shown). This regression remained significant when performed within either of the two largest families (Poaceae and Fabaceae) or the two most speciose genera (*Senecio* and *Vicia*), but no such relationship was observed in two other well-sampled genera (*Vigna* and *Lathyrus*; Table 1). Variation in genome size explained more than 20% of the observed variance in recombination rate in all of the significant regressions except within the Poaceae, where genome size variation explained only ∼8% of the variation in recombination rate.

Phylogenetic estimation is an integral part of nearly all comparative methods (Harvey & Pagel, 1991), but errors in branch length or topology can lead to misleading results (Ackerly, 2000; Symonds, 2002). Parallel

comparative analysis on two independently derived – and incongruent – topologies differing in the treatment of branch lengths provided nearly identical results, suggesting that phylogenetic inaccuracy or error is unlikely to be a serious concern for these data. To further account for the observation that substantial change in both genome size and recombination rate is often evident among species within a genus, comparisons were also made on species-level phylogenies of the two largest families in the data set. All four analyses revealed a similar significant positive correlation between genome size and chiasmata per bivalent (Table 3). Moreover, genome size alone explained a substantial portion of the observed variation in recombination — from more than 10% across all species to nearly 60% of the observed variation in recombination rate within the Fabaceae. For the two largest trees, likelihood ratio tests show that the PGLS model was a significantly better fit than either a nonphylogenetic regression or Felsenstein's (1985) independent contrasts, and the estimated α values suggest that although there is a considerable phylogenetic signal in the correlation, other factors also contribute substantially to the observed variation.

A phylogenetic regression analysis utilizing either of the family-level trees (Fabaceae or Poaceae) revealed a much stronger relationship than that observed for the larger phylogenies. For both of these families, the correlation between genome size and recombination rate was twice that observed in the complete phylogeny, and variation in genome size explained a large portion of the observed variation in recombination rate ($r^2 > 0.35$, Table 3). Estimated values of α for both families were at the extreme high end of the range estimated by the COMPARE software, and log-likelihood values reveal little difference between the PGLS and nonphylogenetic regression models, suggesting little evolutionary constraint acting on the genome size – recombination correlation. Comparison with standard regression results that incorporate data from all taxa in each family suggests that the strength of the above regressions is due at least in part to sampling effects. In the Poaceae, for example, it is likely that exclusion of the genus *Oryza* – with the lowest average genome size in the family, but above average recombination rate – artificially inflated the correlation in the phylogenetic analysis.

To gauge the relative significance of the genome size-recombination rate correlation, I compared the effect of genome size on recombination to the effects of several factors that have previously been linked to differences in recombination. With a reduced data set that included information on perenniality, domestication status, weediness and mating system, I performed a multiple regression of these factors and average bivalent size on recombination rate within the phylogenetic context of the rbcL tree used above (Table 2). Average bivalent size was used instead of total genome size to avoid spurious negative correlations with

Table 1 Nonphylogenetic linear regression of chiasmata per bivalent on genome size and for all available species in each taxon.

Taxon	n	Slope (SE)	r^2
Fabaceae	85	0.14 (0.03)	0.247
Poaceae	63	0.06 (0.03)	0.077
Senecio	30	0.14 (0.04)	0.270
Vicia	19	0.24 (0.09)	0.287
Lathyrus	16	−0.09 (0.09)	0.055
Allium	16	−0.01 (0.09)	0.002

Table 2 Multiple regression of chiasmata per bivalent on bivalent size and several plant life-history characteristics.

	PGLS	FIC	NR
rbcL ($n = 142$)			
α	3.64	0	NA-Large
Log-likelihood	189.73	184.04	167.66
Model r^2	34.1	34.87	38.66
Domesticate	0.06 (0.04)	0.06 (0.04)	0.06 (0.05)
Wild	0.01 (0.04)	0.01 (0.03)	0.06 (0.05)
Weedy	−0.03 (0.04)	−0.04 (0.03)	0.07 (0.04)
Perennial	**−0.06** (0.03)	**−0.06** (0.03)	−0.07 (0.04)
Outcrossing	**−0.12** (0.05)	**−0.14** (0.05)	−0.04 (0.06)
Selfing	0.03 (0.05)	0.00 (0.05)	0.08 (0.06)
Bivalent size	**0.12** (0.02)	**0.11** (0.03)	**0.11** (0.01)
rbcL ($n = 279$)			
α	7.98	0	NA-large
Log-likelihood	301.43	287.83	274.46
Model r^2	13.91	8.27	28.13
Domesticate	0.03 (0.04)	0.03 (0.04)	0.06 (0.04)
Wild	**−0.07** (0.03)	**−0.07** (0.03)	−0.04 (0.04)
Weedy	**−0.07** (0.03)	**−0.08** (0.03)	−0.02 (0.03)
Bivalent size	**0.10** (0.02)	**0.09** (0.02)	**0.11** (0.01)

The partial regression slope and standard error are shown for each characteristic. The three models shown are the phylogenetic generalized least squares (PGLS), Felsenstein's independent contrasts (FIC), and a nonphylogenetic regression (NR). Numbers in bold are significantly different from 0 at the $P < 0.05$ level.

Table 3 Correlation coefficients (r) and slope of regression of chiasmata per bivalent on genome size.

	PGLS	FIC	NR
ITS Fabaceae ($n = 29$)			
α	15.5	0	NA-large
Log-likelihood	30.42	23.27	30.41
r	0.6	0.08	0.74
Slope (SE)	**0.23**(0.06)	0.03 (0.08)	**0.28**(0.05)
ITS Poaceae ($n = 29$)			
α	15.5	0	NA-large
Log-likelihood	40.47	31.31	38.58
r	0.7	0.78	0.62
Slope (SE)	**0.19** (0.04)	**0.19** (0.03)	**0.17**(0.04)
rbcL ($n = 260$)			
α	8.27	0	NA-large
Log-likelihood	276.53	256.93	240.15
r	0.31	0.16	0.46
Slope (SE)	**0.10** (0.02)	**0.07**(0.03)	**0.11** (0.01)
Equal ($n = 279$)			
α	7.5	0	NA-large
Log-likelihood	294.88	281.93	259.48
r	0.3	0.21	0.45
Slope (SE)	**0.10** (0.02)	**0.08** (0.02)	**0.10** (0.01)

The three models shown are the phylogenetic generalized least squares (PGLS), Felsenstein's independent contrasts (FIC), and a nonphylogenetic regression (NR). Numbers in bold are significantly different from 0 at the $P < 0.05$ level.

polyploidy. Results did not differ when total genome size was used and polyploidy included in the analysis or when analyses were performed using the equal-branch length phylogeny. Partial regression results for each character show that genome size has an effect at least as large as the effect of any of the life-history characters surveyed.

Results for the other life-history characters were not entirely congruent with previous results, and should probably be interpreted with caution. Outcrossing taxa were found here to have lower recombination rates than taxa with mixed or selfing mating systems, in agreement both with previous findings (Gibbs *et al.*, 1975; Ross-Ibarra, 2004 and references therein) and theoretical predictions (Roze & Lenormand, 2005). Perennial plants showed lower rates of recombination than annuals or biennials; this contrasts with previous nonphylogenetic results (Ross-Ibarra, 2004), but is not surprising given previous evidence of the relevance of phylogenetic depth to this comparison (Koella, 1993). Weediness showed a significant effect only when analysed on the full tree lacking information on perenniality and mating system; previous workers have presented conflicting evidence on the role of weediness in determining recombination rates (Evans & Weir, 1981; Gornall, 1983) and nonphylogenetic multiple regression has suggested the importance of the interaction between weediness and perenniality (Ross-Ibarra, 2004). Domestication status similarly only

shows a significant effect on the complete tree, although both across-species regression and a more appropriate sister-taxon comparison of domesticates and their wild progenitors using the same data set showed significant results (Ross-Ibarra, 2004).

The general conclusion from the above analyses is that genome size is a significant determinant of recombination rate in Angiosperm species, with an effect as large as that observed for several life-history characteristics. A closer look at genome size and recombination rate data in the genus *Lathyrus*, however, highlights a potentially important caveat in these results: rather than being compared with overall genome size, recombination should be compared with the size of the euchromatic or nonrepetitive fraction of the genome. Heterochromatic or highly repetitive portions of the genome are much less recombinationally active, and empirical work has shown that euchromatin amount is a more important determinant of recombination than total DNA amount among chromosomes within a cell (Sherman & Stack, 1995). Furthermore, amplification of repeat-rich or heterochromatic regions has been implicated as a predominant cause of genome expansion in plants (Bennetzen, 2002). A clear example of the potential problem this causes is evident in a data set of chiasmata and genome size data from eight species of the genus *Lathyrus* (Narayan & McIntyre, 1989). Reanalysis of this data shows that
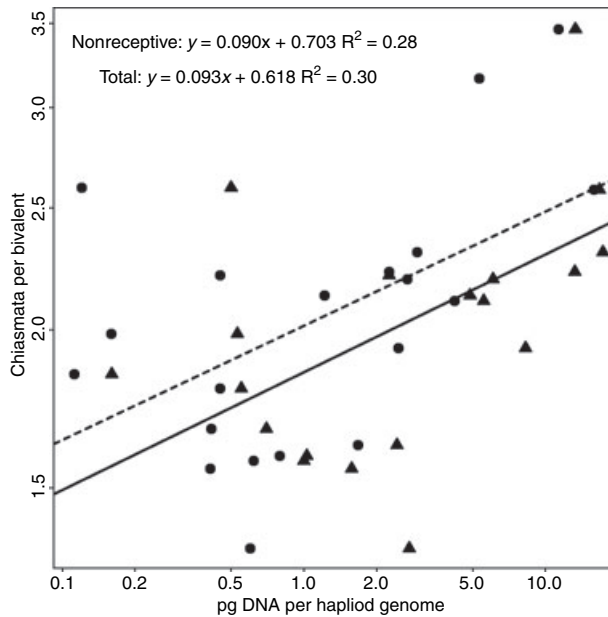
Fig. 2 Linear regression of chiasmata per bivalent on nonrepetitive and total genome size. Circles and the dashed line represent the regression of chiasmata per bivalent on nonrepetitive genome size, whereas triangles and the solid line represent the regression on total genome size. The regression equations shown are for log-transformed data.

although recombination is negatively correlated with total genome size ($r = -0.25$, $P = 0.56$), comparison of recombination to euchromatic genome size reveals a positive correlation ($r = 0.40$, $P = 0.31$). These arguments suggest that the large-scale analyses conducted above could underestimate to a potentially large extent the correlation between euchromatic genome size and recombination. Nonetheless, a comparison of a regression of recombination rate onto size of the nonrepetitive fraction of several plant genomes (Flavell *et al.*, 1974; Wenzel & Hemleben, 1982) with a regression on total genome size revealed no difference in slope (Fig. 2). The data set is small ($n = 18$) and regression results are not corrected for phylogeny, but the analysis suggests that estimates of the effect of genome size on recombination are not likely to be extremely biased.

The analyses presented here show that clear positive correlation between recombination rate and genome size in plant species, and issues of repetitive DNA and genome expansion probably contrive to make the relationship found here an underestimate of the true correlation between genome size and euchromatic DNA. Nonetheless, the slope of the observed log–log regression never exceeded 0.25 and the above survey of the available data on euchromatic genome size provide little reason to believe that the true regression slope would be substantially higher. Recombination events require physical space along a chromosome, and a purely mechanistic

view would predict change in recombination proportional to change in genome size and thus the space available for recombination events. There are several reasons to believe that this expectation is unrealistic, however. All bivalents, regardless of size, must have at least one chiasmata to ensure normal meiotic pairing and assortment (Dawe, 1998), and this effect will be exacerbated if crossover interference extends over a sizable length of the chromosome (Egel, 1995). Moreover, theoretical considerations suggest that even when recombination is selectively favoured a small amount of recombination is usually sufficient to break up correlations between loci (Burt, 2000) and excessive recombination can actually be detrimental (Barton, 1995), such that selection alone could constrain increases in recombination to be less than proportional to change in genome size.

I have shown in this paper that genome size is significantly correlated to recombination rate across a wide sampling of Angiosperm species even after correction for phylogenetic effects, and that change in genome size explains a meaningful proportion of the observed variation in recombination rate. Although the magnitude of the observed effect of genome size on recombination rate is comparable with the effects of several life-history characters previously linked to evolutionary change in recombination rate, consideration of processes of genome size change in angiosperms suggests that the observed correlation is likely a conservative estimate of the relationship between recombination and euchromatic genome size. And finally, although recombination rate was found to increase less than proportionally to change in genome size, a number of theoretical and mechanistic arguments suggest that this result is not unexpected.

Acknowledgements

References

Ackerly, D.D. 2000. Taxon sampling, correlated evolution, and independent contrasts. *Evolution* **54**: 1480–1492.

Anderson, L.K., Hooker, K.D. & Stack, S.M. 2001. The distribution of early recombination nodules on zygotene bivalents from plants. *Genetics* **159**: 1259–1269.

Anderson, L., Doyle, G., Brigham, B., Carter, J., Hooker, K.D., Lai, A., Rice, M. & Stack, S.M. 2003. High-resolution crossover maps for each bivalent of *Zea mays* using recombination nodules. *Genetics* **165**: 849–865.

Angiosperm Phylogeny Group 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Bot. J. Linn. Soc.* **141**: 399–436.

Barton, N.H. 1995. A general model for the evolution of recombination. *Genet. Res.* **65**: 123–145.

Bennett, M. & Leitch, I. 2005). *Plant DNA C-values Database*. V. 4.0. URL http://www.kew.org/cval/homepage.html.

Bennetzen, J.L. 2002. Mechanisms and rates of genome expansion and contraction in flowering plants. *Genetica* **115**: 29–36.

Burt, A. 2000. Perspective: sex, recombination, and the efficacy of selection – was Weismann right? *Evolution* **54**: 337–351.

Cavalier-Smith, T. 1985. Eukaryote gene numbers, non-coding DNA and genome size. In: *The Evolution of Genome Size* (Cavalier-Smith, T., ed.), pp. 69–104. John Wiley and Sons Ltd, London.

Dawe, R.K. 1998. Meiotic chromosome organization and segregation in plants. *Annu. Rev. Plant. Physiol. Plant Mol. Biol.* **49**: 371–395.

Downie, S., Katz-Downie, D. & Watson, M. 2000. A phylogeny of the flowering plant family Apiaceae based on chloroplast DNA rpl16 and rpoC1 intron sequences: towards a suprageneric classification of subfamily Apioideae. *Am. J. Bot.* **87**: 273–292.

Egel, R. 1995. The synaptonemal complex and the distribution of meiotic recombination events. *Trends Genet.* **11**: 206–208.

Evans, A. & Weir, J. 1981. The evolution of weed beet in sugar beet crops. *Genet. Resour. Crop Evol.* **29**: 301–310.

Felsenstein, J. 1985. Phylogenies and the comparative method. *Am. Nat.* **125**: 1–15.

Felsenstein, J. 1993). *Phylip (Phylogeny Inference Package)* version 3.6. distributed by the author. Department of Genome Sciences, University of Washington, Seattle, WA.

Flavell, R., Bennet, M., Smith, J. & Smith, D. 1974. Genome size and proportion of repeated nucleotide sequence DNA in plants. *Biochem. Genet.* **12**: 257–259.

Gibbs, P., Milne, C. & Carillo, M. 1975. Correlation between the breeding system and recombination index in five species of *Senecio*. *New Phytol.* **75**: 619–626.

Gornall, R.J. 1983. Recombination systems and plant domestication. *Biol. J. Linn. Soc.* **20**: 375–383.

Harinarayana, G. & Murty, B.R. 1971. Cytological regulation of recombination in *Pennisetum* and *Brassica*. *Cytologia* **36**: 435–448.

Harvey, P.H. & Pagel, M.D.P. 1991). *The Comparative Method in Evolutionary Biology*. Oxford University Press, Oxford.

Hsiao, C., Chatterton, N.J., Asay, K.H. & Jensen, K.B. 1995. Phylogenetic relationships of the monogenomic species of the wheat tribe, Triticeae (Poaceae), inferred from nuclear rDNA (internal transcribed spacer) sequences. *Genome* **38**: 211–223.

Hu, J.M., Lavin, M., Wojciechowski, M.F. & Sanderson, M.J. 2000. Phylogenetic systematics of the tribe Millettieae (Leguminosae) based on chloroplast trnK/matK sequences and its implications for evolutionary patterns in Papilionoideae. *Am. J. Bot.* **87**: 418–430.

Jakob, S.S., Meister, A. & Blattner, F.R. 2004. The considerable genome size variation of *Hordeum* species (Poaceae) is linked to phylogeny, life form, ecology, and speciation rates. *Mol. Biol. Evol.* **21**: 860–9.

Kajita, T., Ohashi, H., Tateishi, Y., Bailey, C.D. & Doyle, J.J. 2001. rbcL and legume phylogeny, with particular reference to Phaseoleae, Millettieae, and allies. *Syst. Bot.* **26**: 515–536.

Koella, J. 1993. Ecological correlates of chiasma frequency and recombination index of plants. *Biol. J. Linn. Soc.* **48**: 227–238.

Levin, D. & Funderburg, S. 1979. Genome size in angiosperms: temperate vs. tropical species. *Am. Nat.* **114**: 784–795.

Martins, E. 2004). *COMPARE. Computer Programs for the Statistical Analysis of Comparative Data*. Distributed by the author at http://compare.bio.indiana.edu/. Department of Biology, Indiana University, Bloomington, IN.

Martins, E.P. & Hansen, T. 1997. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into analysis of interspecific data. *Am. Nat.* **149**: 646–667.

Mason-Gamer, R.J., Orme, N.L. & Anderson, C.M. 2002. Phylogenetic analysis of North American *Elymus* and the monogenomic Triticeae (Poaceae) using three chloroplast DNA data sets. *Genome* **45**: 991–1002.

Narayan, R. & McIntyre, F.K. 1989. Chromosomal DNA variation, genomic constraints and recombination in *Lathyrus*. *Genetica* **79**: 45–52.

Olmstead, R., Sweere, J., Spangler, R. & Jd, J.P. 1999. Phylogeny and provisional classification of the Solanaceae based on chloroplast DNA. In: *Solanaceae IV* (Nee, M., Lester, R. & Hawkes, J., eds), pp. 111–137. Royal Botanic Gardens, Kew.

Rees, H. & Dale, P.J. 1974. Chiasmata and variability in *Lolium* and *Festuca* populations. *Chromosoma* **47**: 335–351.

Rees, H. & Durrant, A. 1986. Recombination and genome size. *Theor. Appl. Genet.* **73**: 72–76.

Ross-Ibarra, J. 2004. The evolution of recombination under domestication: a test of two hypotheses. *Am. Nat.* **163**: 105–112.

Roze, D. & Lenormand, T. 2005. Self-fertilization and the evolution of recombination. *Genetics* **170**: 841–857.

Salamin, N., Hodkinson, T.R. & Savolainen, V. 2002. Building supertrees: an empirical assessment using the grass family (Poaceae). *Syst. Biol.* **51**: 136–150.

Sherman, J.D. & Stack, S.M. 1995. Two-dimensional spreads of synaptonemal complexes from Solanaceous plants. VI. High-resolution recombination nodule map for tomato (*Lycopersicon esculentum*). *Genetics* **141**: 683–708.

Soltis, D.E., Soltis, P.S., Bennett, M.D. & Leitch, I.J. 2003. Evolution of genome size in the angiosperms. *Am. J. Bot.* **90**: 1596–1603.

Steele, K. & Wojciechowski, M. 2003. Phylogenetic analyses of tribes Trifolieae and Vicieae based on sequences of the plastid gene matK (Papilionoideae: Leguminosae). In: *Advances in Legume Systematics*, Part 10 (B. Klitgaard & A. Bruneau eds), pp. 355–370. Royal Botanic Gardens, Kew.

Stevens, P. 2005). *Angiosperm Phylogeny Website*. URL http://www.mobot.org/MOBOT/research/APweb/

Symonds, M.R.E. 2002. The effects of topological inaccuracy in evolutionary trees on the phylogenetic comparative method of independent contrasts. *Syst. Biol.* **51**: 541–553.

Tenaillon, M.I., Sawkins, M.C., Anderson, L.K., Stack, S.M., Doebley, J. & Gaut, B.S. 2002. Patterns of diversity and recombination along chromosome 1 of maize (*Zea mays* ssp. *mays* L.). *Genetics* **162**: 1401–1413.

Thompson, J., Higgins, D. & Gibson, T. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap

penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673–4680.

Wenzel, W. & Hemleben, V. 1982. A comparative study of genomes in angiosperms. *Plant Sys. Evol.* **139**: 209–227.

Supplementary Material

The following supplementary material is available for this article online:

Figure S1. Phylogenies used for analysis. A. Equal branch lengths. B. Chloroplast rbcL neighbour-joining tree. C. Fabaceae rDNA neighbour-joining tree. D. Poaceae rDNA neighbour-joining tree. In A and B species within a genus are coded as polytomies (not shown).

This material is available as part of the online article from http://www.blackwell-synergy.com