doi:10.1111/j.1558-5646.2010.01045.x

AN INTEGRATIVE TEST OF THE DEAD-END HYPOTHESIS OF SELFING EVOLUTION IN TRITICEAE (POACEAE)

Juan S. Escobar,[1,2,3] Alberto Cenci,[4,5] Jeremy Bolognini,[1] Annabelle Haudry,[6] Stefan Laurent,[2,7] Jacques David,[1,4] and Sylvain Glémin[2]

[1]Institut National de la Recherche Agronomique, Centre de Montpellier, UMR Diversité et Adaptation des Plantes Cultivées, Domaine de Melgueil, 34130 Mauguio, France

[2]Institut des Sciences de l'Evolution, UMR 5554, Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France

[3]E-mail: juan-sebastian.escobar@univ-montp2.fr

[4]Montpellier Supagro, Centre International d'Etudes Supérieures en Sciences Agronomiques, UMR Diversité et Adaptation des Plantes Cultivées, 2 Place Pierre Viala, 34060 Montpellier Cedex 1, France

[6]Division of Ecology and Evolutionary Biology, Graham Kerr Building, University of Glasgow, Glasgow G12 8QQ, Scotland, United Kingdom

Received December 2, 2009
Accepted April 21, 2010

Self-fertilization is hypothesized to be an evolutionary dead end because reversion to outcrossing can rarely happen, and selfing lineages are thought to rapidly become extinct because of limited potential for adaptation and/or accumulation of deleterious mutations. We tested these two assumptions by combining morphological characters and molecular-evolution analyses in a tribe of hermaphroditic grasses (Triticeae). First, we determined the mating system of the 19 studied species. Then, we sequenced 27 protein-coding loci and compared base composition and substitution patterns between selfers and outcrossers. We found that the evolution of the mating system is best described by a model including outcrossing-to-selfing transitions only. At the molecular level, we showed that regions of low recombination exhibit signatures of relaxed selection. However, we did not detect any evidence of accumulation of nonsynonymous substitutions in selfers compared to outcrossers. Additionally, we tested for the potential deleterious effects of GC-biased gene conversion in outcrossing species. We found that recombination and not the mating system affected substitution patterns and base composition. We suggest that, in Triticeae, although recombination patterns have remained stable, selfing lineages are of recent origin and inbreeding may have persisted for insufficient time for differences between the two mating systems to evolve.

KEY WORDS: Biased gene conversion, effective population size, mating system, protein evolution, recombination, selection efficiency, substitution rate.

[5]Present address: Institut de Recherche pour le Développement, UMR RPB-Equipe DIVA, 911 Avenue Agropolis, 34394 Montpellier Cedex 5, France.

[7]Present address: Section of Evolutionary Biology, Ludwig Maximilians Universität BioCenter, Grosshaderner Str. 2, 82152 Planegg-Martinsried, Germany.

Since Darwin (1876, 1878), the astonishing diversity of mating systems observed in plants has generated an impressive amount of theoretical and empirical work to understand the distribution of such systems across species and the causes of their evolution. A commonly admitted idea has been popularized by Stebbins (1957), who suggested that self-fertilization should be an evolutionary dead end. Selfing lineages would continually become extinct and unilateral transitions from outcrossing to selfing would recurrently found new lineages. On the short term, self-fertilization could get fixed because it has two main advantages over outcrossing: reproductive assurance under pollen limitation (Baker 1955, 1967) and the twofold transmission of genes (Fisher 1941). Inbreeding depression, that is, the reduced fitness of selfed relative to outbred offspring (Lande and Schemske 1985; Charlesworth and Charlesworth 1987), counteracts the advantages of selfing and explains the maintenance of outcrossing (Goodwillie et al. 2005; Charlesworth 2006). Theoretical models predict that selfing populations should purge their inbreeding depression (Lande and Schemske 1985) such that, once self-fertilization has evolved, reversion to outcrossing is not possible because inbreeding depression is too low to overcome the advantages of selfing.

Despite its short-term advantages, selfing is thought to have negative long-term evolutionary consequences because it strongly affects population characteristics. First, selfing is expected to lead to an automatic reduction in the effective population size by reducing the number of independent gametes sampled for reproduction (Pollak 1987; Nordborg 2000). Second, founding effects are expected to be more frequent and more severe in selfers because a single seed can found a new population (Baker 1955, 1967), which may strongly reduce the effective population size. Strong extinction–recolonization dynamics should also reduce effective population size in selfers (Ingvarsson 2002). Third, selfers suffer from reduced effective recombination compared to outcrossers due to the reduced heterozygosity, which leads to increased hitchhiking, such as selective sweeps (Kaplan et al. 1989) and background selection (Charlesworth et al. 1993, 1995), subsequently reducing the effective population size (Nordborg et al. 1996; Charlesworth and Wright 2001; Charlesworth 2009). As an overall consequence, genetic diversity should be lower in selfers than in outcrossers. Moreover, because of both reduced effective population size and effective recombination rate, selection efficiency is expected to be lower in selfers than in outcrossers (Charlesworth 2009). Thus, selfing species should be less efficient at eliminating slightly deleterious alleles or fixing new advantageous mutations than outcrossing species (Charlesworth 1992; Glémin 2007). Consequently, selfing lineages would eventually become extinct because of limited adaptive potential and/or accumulation of slightly deleterious mutations (Lynch et al. 1995).

Irreversibility of mating-system transitions and maladaptation in selfing lineages constitute the two assumptions of the dead-end hypothesis of selfing evolution (Stebbins 1957; reviewed in Takebayashi and Morrell 2001). Empirical studies in *Dalechampia* (Armbruster 1993), Polemoniaceae (Barrett et al. 1996b), Ponterediaceae (Kohn et al. 1996), *Amsinckia* (Schoen et al. 1997), *Linanthus* (Goodwillie 1999) and Solanaceae (Igic et al. 2006) support the assumption that selfing lineages evolve from outcrossing ancestors in flowering plants. For instance, it has been shown that the breakdown of self-incompatibility systems is much more frequent than the gain of self-incompatibility (Goodwillie 1999; Igic et al. 2006). To date, few studies have tested the second assumption of the dead-end hypothesis of selfing evolution. Studies using diverse genetic markers (microsatellites, RAPD, allozymes and sequences) have shown that selfers have lower genetic diversity than outcrossers (Hamrick and Godt 1996; Nybom 2004; Glémin et al. 2006), although evidence is limited for quantitative-genetic variation (Charlesworth and Charlesworth 1995). Direct evidence of lack of adaptive potential in selfers is much more difficult to obtain. However, the accumulation of deleterious mutations is testable through the analysis of substitution patterns in protein-coding sequences. According to the dead-end hypothesis, we expect to observe signatures of relaxed selection in selfers, such as an elevated ratio of nonsynonymous to synonymous substitution rates (ω) or weak codon usage bias. Studies in few *Arabidopsis*, *Caenorhabditis* and Triticeae grass species give little support to changes in nucleotide substitution patterns between selfing and outcrossing species (Wright et al. 2002; Cutter et al. 2008; Haudry et al. 2008). One hypothesis put forward by authors to explain results in *Arabidopsis* and *Caenorhabditis* is the recent origin of selfing in these two genera. An alternative explanation was proposed by Haudry et al. (2008) to explain results in grasses: outcrossing species could pay the cost of a load associated to the GC-biased gene conversion, although that cost should be very low, if any, in selfing species.

GC-biased gene conversion (hereafter gBGC) is a segregation distortion associated to recombination favoring G and C over A and T alleles (reviewed in Marais 2003; Duret and Galtier 2009). gBGC is increasingly recognized as a major force structuring genomes in mammals (Galtier et al. 2001; Duret and Arndt 2008), birds (Webster et al. 2006), yeast (Birdsell 2002), and it likely occurs in grasses (Glémin et al. 2006; Haudry et al. 2008) and maybe in other plants (Wright et al. 2006). gBGC has direct impact on the GC content. In addition, because it increases the fixation probability of G and C bases it may also oppose natural selection by promoting the fixation of deleterious AT $\rightarrow$ GC mutations. In highly recombining regions, gBGC could give spurious signatures of relaxed or even positive selection by increasing ω ratios (Galtier and Duret 2007; Galtier et al. 2009). Selfing species are expected to have reduced gBGC efficiency because of their

low heterozygosity. Thus, selfing species should exhibit lower GC content than outcrossing ones (Marais et al. 2004; Wright et al. 2007; Haudry et al. 2008). In addition, outcrossers could pay a new genomic load associated with gBGC, through the accumulation of deleterious G or C mutations (the so-called "genomic Achilles' heel," Galtier and Duret 2007; Galtier et al. 2009). This challenges one of the basic assumptions of the dead-end hypothesis of selfing evolution: rather than being a straightforward consequence of the mating system, substitution patterns and genome composition could depend on the balance between the reduction in selection efficacy expected in selfers and the cost of recombination associated to gBGC in outcrossers.

To our knowledge, no study has tested simultaneously the two assumptions of the dead-end hypothesis of selfing evolution so far. In addition, if the gBGC is active in grasses, its cost in terms of genetic load in outcrossing species needs to be evaluated. In this article, we use 19 diploid hermaphroditic Triticeae grasses to test the two assumptions of the dead-end hypothesis and the cost of the gBGC. We use different morphological characters to determine the preferential mating system of the studied species and test assumptions of the dead-end hypothesis of selfing evolution. For this, we use a recent multigenic phylogeny of the tribe to infer whether transitions between the two mating systems are uni- or bidirectional. At the molecular level, we use a set of 27 genes to test whether substitution patterns and genome composition differ between selfers and outcrossers. We also took advantage of the chromosomal location of these genes to perform parallel analyses on the effect of recombination on molecular patterns.

Materials and Methods

STUDIED SPECIES

Triticeae is a tribe within the Pooideae subfamily of grasses including species of major economic importance, like wheat, barley, and rye. The tribe comprises annual and perennial, wind-pollinated species. We obtained morphological and molecular data in 19 diploid species, spanning 13 genera. These species were selected because they belong to most phylogenetic clades recognized so far (Kellogg et al. 1996; Petersen and Seberg 1997; Mason-Gamer 2005) and represent most of the diversity of diploid genera (68% according to Kellogg et al. 1996 and Seberg and Frederiksen 2001), life styles (annual and perennial), mating systems (self-compatible and self-incompatible), and geographical location (Europe, Middle East, Asia, North America and Australia) of the tribe. In addition, they were easily obtained from the National Plant Germplasm System of the US Department of Agriculture (http://www.ars-grin.gov/npgs/index.html). One or two accessions per species were obtained, making a total of 31 accessions (Table S1). In addition, *Brachypodium* sp. and *Brachypodium distachyon* were used as outgroups in all comparative analyses. Sequences of *B. distachyon* were obtained from the US Department of Energy, Joint Genome Institute (http://www.jgi.doe.gov/).

MATING-SYSTEM DETERMINATION

Triticeae are known to have diverse mating systems, spanning from obligate cross-fertilization due to strict self-incompatibility (e.g., rye—*Secale cereale*) to strong self-fertilization (e.g., barley—*Hordeum vulgare* and diploid wheat—*Triticum monococcum*). They are thus an ideal target to study evolutionary transitions between selfing and outcrossing. However, excepting some well-studied species, the mating system of most of the Triticeae remains unknown. In the absence of selfing-rate estimates for each of the 19 studied species, we determined the mating system using three traits thought to accurately describe it: the autonomous seed set, pollen/ovule ratio, and anther size (Cruden 1977, 2000; Schoen 1982; Damgaard and Loeschcke 1994). For this, one to three plants per accession were individually sown in October–November 2005 in 3-l plastic recipients and randomized in the glass house. To estimate the self-fertilization capacity of each species, approximately half of spikes of each plant were surrounded with plastic bags, which effectively impeded cross-fertilization. At the end of the reproductive season (March–April 2006), we counted the number of self-fertilized seeds and the number of spikelets per spike. We also removed three spikelets and six mature anthers per plant for morphological measures (see below).

The autonomous seed set was estimated as:

$$\frac{N_{\text{self-fertilized seeds}}}{N_{\text{self-fertilized spikelets}} \times N_{\text{fertile flowers/spikelet}}}. \quad (1)$$

In Triticeae, spikelets bear generally three flowers, excepting some species that can bear five to six flowers. In the majority of species, the central flower is sterile and the two lateral flowers are fertile, although the inverse pattern is sometimes observed (e.g., *Hordeum*). We noted the origin of each seed obtained in self-fertilized spikelets and determined the number of fertile flowers per spikelet in each accession and species.

To estimate the flower size, we photographed the removed spikelets of each plant on graph paper. Using digital images, we measured glumes and lemmas (i.e., basal inflorescence bracts) with ImageJ 1.34s (Rasband 2007), and estimated the flower size as the mid value between these two parts of the flower. In some species (e.g., *H. vulgare*) glumes or lemmas bear bristles (awns); awns were neglected in the flower size estimation. The removed anthers were photographed under the stereoscope and the anther size was determined through the analysis of digital images with ImageJ 1.34s. The number of pollen grains was counted in a particle counting analyzer (Multisizer 3, Beckman

Coulter Inc., Fullerton, CA) from six anthers removed in each accession and species. Because only one ovule per flower is produced in Triticeae, the number of pollen grains directly estimates the pollen/ovule ratio, an accurate measure of the mating system, especially reliable in wind-pollinated species (Cruden 1977; Michalski and Durka 2009). Pollen/ovule ratio and anther size were scaled to the flower size to take into account allometric effects.

We correlated the autonomous seed set with the pollen/ovule ratio (transformed using the natural logarithm) and anther size (both measures scaled to the flower size), and classified species as selfing or outcrossing. The distribution of the mating system has been shown to be strongly bimodal in wind-pollinated species (Schemske and Lande 1985; Goodwillie et al. 2005). We thus used a simple dichotomic classification to simplify analyses. However, this classification might obscure patterns of mating-system evolution if mixed-mating species are pooled with actual selfing and outcrossing species. Intermediate or undetermined selfing rates were taken into account in molecular-evolution analyses (see below). Species with known mating systems were used to guide our classification. These include one perennial outcrosser (*Psathyrostachys juncea*, selfing rate (S) = 0; Yang et al. 2008), three annual outcrossers (*Dasypyrum villosum*, $S = 0.25$, Depace and Qualset 1995; *S. cereale*, $S = 0$; Polanco et al. 1994; and *Aegilops speltoides*, unknown selfing rate but reported as allogamous; Zohary and Imber 1963; Dvorak et al. 1998; Zaharieva and Monneveux 2006), and four annual selfers (*Ae. tauschii*, $S = 0.98$; Dvorak et al. 1998; *T. monococcum*, $S = 0.95$; Hegde et al. 2000; *H. vulgare*, $S = 0.98$; Kahler et al. 1975; von Bothmer et al. 1995; Parzies et al. 2000; Abdel-Ghani et al. 2004; Morrell et al. 2005; and *H. marinum*, unknown selfing rate but reported as inbred; von Bothmer et al. 1995).

SAMPLED LOCI

We used 27 orthologous nuclear loci, located on four different chromosomes of the seven chromosomes representative of Triticeae, previously used to reconstruct a multigenic phylogeny of Triticeae (J. S. Escobar et al., unpubl. ms.). Details of loci and sequencing protocols are presented elsewhere (J. S. Escobar et al., unpubl. ms.). Here, we summarize the key points. From the 27 sequenced loci, 21 were derived from the rice chromosome 1, known to be collinear to the Triticeae chromosome 3 (Sorrells et al. 2003; Munkvold et al. 2004; Haudry et al. 2008). As the wheat genome is not sequenced nor assembled yet, we used the location of rice orthologs as a proxy of their chromosomal position in Triticeae. We verified that using the draft sequence of *B. dystachion* as reference did not alter physical position. Additionally, orthologs of one gene fragment corresponding to a eukaryotic initiation factor involved in translational regulation (*eIFiso4E*), located on the long chromosome 1 arm; three tightly linked loci, corresponding to the hardness gene (*GSP*, *PinA* and *PinB*; Chantret et al. 2005), located in telomeric position on the short chromosome 5 arm; and two gene fragments involved in the carotenoid biosynthetic pathway (*CRTISO* located on the long chromosome 4 arm, and *PSY2* in chromosome 5; Sorrells et al. 2003; Cenci et al. 2004), were sequenced (Table 1).

Raw sequence data were aligned with the Staden Package (Staden et al. 2000) and resulting alignments were manually corrected. When two accessions per species were available, the consensus sequence has been built and used for the analyses with BioEdit 7.0 (Hall 1999). This is because we were mainly interested in the way that mating systems affect molecular divergence among species, that is, the pattern of fixation of variants arising by mutation, not segregating polymorphisms for which we do not have the appropriate dataset. Using one sequence per species rather than the consensus do not change results.

PHYLOGENETIC ANALYSES

Detailed methods for the phylogenetic reconstruction of Triticeae are presented elsewhere (J. S. Escobar et al., unpubl. ms.). Briefly, analyses were performed on individual loci and the concatenate of all loci using maximum likelihood (ML) and Bayesian approaches. ML analyses were conducted using the best-fitting model of sequence evolution. Model selection was based on Akaike's information criterion (AIC) using ModelTest 3.7 (Posada and Crandall 1998). PAUP* 4.0b10 (Swofford 2003) was used to obtain the log-likelihood and the phylogenetic trees (heuristic search with neighbor-joining starting tree, tree bissection-reconnection swapping, and 100 bootstrap replicates). Bayesian analyses were performed with MrBayes 3.1.2 (Huelsenbeck and Ronquist 2001). Markov chain Monte Carlo (MCMC) analyses were run with random starting trees and five simultaneous, sequentially heated independent chains. Analyses were run during 1,000,000 generations for individual loci and 10,000,000 generations for the concatenated loci. We used the BPCOMP program implemented in PhyloBayes 2.3c (Lartillot and Philippe 2004) to determine appropriate convergence of the chains. A burn-in was established after identifying the stationary phase.

Individual gene trees (Table S2) are sometimes incongruent with the tree of all concatenated loci (Fig. 2) (J. S. Escobar et al., unpubl. ms.). Incongruence was quantified using a homemade script comparing tree topologies. For this, we obtained the consensus tree of the 100 bootstrap trees of each locus and unresolved all nodes with less than 70% support. This permitted to distinguish actual incongruence from stochastic errors due to low-resolution power of individual loci. We counted the number of triplets of species in the gene tree and the concatenated tree (restricted to the number of species of each gene), and calculated the percentage of triplets of species that were present in the gene

Table 1. Genomic location, GC composition, relevant substitution parameters, and incongruence with the multigenic tree for all sequenced loci. Chr., chromosome; S, short arm; L, long arm; Tel., telomere; Cen., centromere; NA, not available; RDC, relative distance to the centromere; GC_{123}, GC content at all codon positions; GC_n, GC content at the *n*th codon position; d_N, nonsynonymous substitution rate; d_S, synonymous substitution rate; Incongr., proportion of incongruent triplets between the gene tree and the multigenic tree. GC* and ω were estimated using the phylogeny proper to each locus. Loci are sorted by increasing RDC.

Locus	Location	RDC	Size (bp)	Incongr.	GC_{123}	GC_1	GC_2	GC_3	GC_{123}*	GC_3*	ω	d_N	d_S
LOC_Os01g37560	Chr. 3L, Cen.	0.1598	1005	0.027	0.566	0.595	0.441	0.662	0.4535	0.4581	0.1446	0.0974	0.7217
LOC_Os01g24680	Chr. 3S, Cen.	0.1841	1014	0.017	0.442	0.529	0.377	0.419	0.4534	0.4153	0.0722	0.0515	0.7199
LOC_Os01g39310	Chr. 3L, Cen.	0.2017	945	0.047	0.472	0.541	0.445	0.430	0.3924	0.3200	0.0769	0.0643	0.7092
LOC_Os01g21160	Chr. 3S, Cen.	0.3067	1017	0.032	0.446	0.554	0.369	0.417	0.5397	0.5187	0.2422	0.2431	0.9878
LOC_Os01g19470	Chr. 3S, Cen.	0.3516	942	0.016	0.424	0.530	0.440	0.302	0.3936	0.2745	0.2552	0.1480	0.6442
LOC_Os01g48720	Chr. 3L, Cen.	0.4172	939	0.052	0.413	0.520	0.374	0.346	0.5268	0.4888	0.4064	0.2850	0.6273
LOC_Os01g53720	Chr. 3L, Cen.	0.5261	1101	0.022	0.417	0.525	0.367	0.359	0.3862	0.3346	0.2357	0.1453	0.5913
LOC_Os01g55530	Chr. 3L, Cen.	0.5674	1068	0.021	0.489	0.542	0.454	0.467	0.5731	0.5803	0.0954	0.0738	0.8047
LOC_Os01g13200	Chr. 3S, Cen.	0.5677	897	0.012	0.438	0.470	0.389	0.457	0.3825	0.3824	0.0514	0.0325	0.7579
LOC_Os01g56630	Chr. 3L, Cen.	0.5925	915	0.018	0.424	0.498	0.314	0.462	0.3616	0.3808	0.1514	0.0751	0.5271
LOC_Os01g11070	Chr. 3S, Cen.	0.6525	1050	0.251	0.443	0.485	0.409	0.436	0.3754	0.3418	0.1504	0.1043	0.6830
LOC_Os01g60230	Chr. 3L, Cen.	0.6726	999	0.038	0.449	0.548	0.379	0.421	0.4787	0.4799	0.0806	0.0663	0.6108
LOC_Os01g61720	Chr. 3L, Tel.	0.7050	933	0.025	0.520	0.581	0.489	0.490	0.4205	0.3878	0.0789	0.0950	0.7157
LOC_Os01g09300	Chr. 3S, Tel.	0.7221	861	0.078	0.461	0.572	0.364	0.448	0.4731	0.5078	0.1066	0.0834	0.7621
LOC_Os01g62900	Chr. 3L, Tel.	0.7318	951	0.061	0.417	0.563	0.360	0.328	0.5565	0.5237	0.0483	0.0382	0.7887
LOC_Os01g67220	Chr. 3L, Tel.	0.8271	1101	0.116	0.413	0.522	0.363	0.356	0.4985	0.4453	0.1596	0.1306	0.6929
LOC_Os01g68770	Chr. 3L, Tel.	0.8620	996	0.082	0.521	0.589	0.415	0.556	0.7003	0.7277	0.0508	0.0582	1.1820
LOC_Os01g70670	Chr. 3L, Tel.	0.8979	882	0.017	0.455	0.578	0.471	0.318	0.4537	0.3230	0.1024	0.0758	0.7538
LOC_Os01g72220	Chr. 3L, Tel.	0.9334	1131	0.233	0.480	0.602	0.460	0.376	0.4078	0.2914	0.0925	0.0639	0.5874
LOC_Os01g73790	Chr. 3L, Tel.	0.9654	966	0.000	0.551	0.561	0.378	0.713	0.7556	0.8032	0.0472	0.0420	0.8708
LOC_Os01g01790	Chr. 3S, Tel.	0.9757	858	0.025	0.466	0.575	0.397	0.428	0.6402	0.5968	0.0246	0.0536	1.2457
eIFiso4E	Chr. 1L	NA	627	0.121	0.515	0.526	0.427	0.591	0.5935	0.5771	0.1056	0.0666	0.5700
CRTISO	Chr. 4L	NA	525	0.029	0.418	0.479	0.330	0.445	0.4648	0.4550	0.0839	0.0845	0.8845
GSP	Chr. 5S, Tel.	NA	492	0.584	0.510	0.500	0.493	0.539	0.3520	0.3896	0.3631	0.3155	0.8696
PinA	Chr. 5S, Tel.	NA	450	0.000	0.547	0.552	0.500	0.586	0.5357	0.4972	0.5020	0.2253	0.4084
PinB	Chr. 5S, Tel.	NA	450	0.192	0.531	0.526	0.477	0.590	0.2676	0.2292	0.2557	0.1920	0.7560
PSY2	Chr. 5	NA	459	0.000	0.530	0.530	0.604	0.555	0.6202	0.6845	0.0042	0.0032	0.7603

tree and absent in the concatenated tree. Loci with less than 5% of incongruence were considered congruent, otherwise incongruent (Table 1).

EVOLUTIONARY TRANSITIONS IN MATING SYSTEMS

Mating systems were mapped onto the Triticeae phylogeny reconstructed with the 27 loci (Fig. 2). Transitions between outcrossing and selfing were determined by parsimony using the reconstruction of ancestral states package implemented in Mesquite 2.5 (Maddison and Maddison 2008). In addition, ML and MCMC procedures were applied to the 100 bootstrap trees from which we obtained support values of nodes. Analyses of bootstrap trees allowed assessing the uncertainty of transitions in nodes not fully supported. ML and MCMC analyses on bootstrap trees were performed with the BayesMultistates program (Pagel et al. 2004) implemented in BayesTraits 1.0. MCMC analyses were run during 5,050,000 generations, with *ratedev* parameter = 100 (this parameter specifies how big a change is proposed to the rate coefficients at each iteration of the Markov chain), a uniform prior distribution and a burn-in of 50,000 generations.

Three different models were compared using likelihood ratio tests: (1) the unrestricted model in which the probability of the two types of transitions, from outcrossing to selfing (q_{os}) and from selfing to outcrossing (q_{so}), were calculated; (2) a restricted model in which only outcrossing to selfing transitions were permitted (i.e., $q_{so} = 0$); and (3) an alternative, restricted model in which only selfing to outcrossing transitions were permitted (i.e., $q_{os} = 0$). The unrestricted model estimates three parameters: the two transition probabilities (q_{os} and q_{so}) and the probability that the mating system at the root of the tree was outcrossing (q_o; $q_s = 1 - q_o$). In the restricted models, ancestral states were fixed to outcrossing or selfing, respectively.

MOLECULAR-EVOLUTION ANALYSES

Aligned sequences were analyzed in two ways: (1) analyses of individual loci and (2) analyses of concatenated loci. Individual locus alignments were first analyzed with MEGA 4 (Tamura et al. 2007) to determine base composition and GC content. The mean and median GC content across all loci for the total sequence and for each of the three codon positions were calculated. We then analyzed substitution rates and estimated the ratio of nonsynonymous (d_N) to synonymous (d_S) substitution rates (ω), and the equilibrium GC content (or GC*). GC* is defined as

$$GC^* = \frac{AT \rightarrow GC}{AT \rightarrow GC + GC \rightarrow AT}, \qquad (2)$$

where AT → GC refers to the number of substitutions from A or T to G or C bases, and GC → AT holds for the inverse (Sueoka 1962). GC* ratios were calculated at all codon positions (GC_{123}*) and at the third codon position (GC_3*). We used a maximum-likelihood approach to estimate substitution ratios (ω, GC_{123}*, and GC_3*) across branches of phylogenies under various hierarchical models of sequence evolution. These models were fitted using the CODEML program implemented in the PAML 4.1 package (Yang 2007) and the BPPML 0.3.1 program implemented in the Bio++ suite of libraries and programs (Dutheil and Boussau 2008). We used likelihood-ratio tests to assess whether more complex models provided a significantly improved fit compared to simpler models.

Positive selection was tested on individual loci by comparing the nearly neutral (M1a) and positive selection (M2a) models (Yang et al. 2005). The M1a model allows two categories of sites: sites with $\omega < 1$ and sites with $\omega = 1$. The M2a model is similar to model M1a but allows a third category of sites ($\omega > 1$). In addition, four hierarchical, nested models were compared. (1) The one-ratio model constrained all branches of the phylogeny to have the same substitution rate (Fig. S1A). (2) The two-ratio model estimated one substitution ratio for the outgroup and another one for the rest of the branches (i.e., Triticeae branches; Fig. S1B). (3) The three-ratio model estimated one ratio for the outgroup, a second ratio for internal branches, and a third ratio for external branches. Internal branches were those for which the mating system could not be inferred (e.g., branches connecting selfing and outcrossing species); external branches were those for which the mating system could be inferred (i.e., branches for which the mating system was known, was it selfing or outcrossing; Fig. S1C). (4) The four-ratio model was similar to model (3), but in this model we estimated different ratios for external selfing and outcrossing branches. Note that branches connecting two or more selfing or outcrossing species were considered as selfing or outcrossing, respectively, excepting the branch leading to *Hordeum*, which was laid as undetermined because the existence of self-incompatible species within the genus was (Baumann et al. 2000; Blattner 2004), not analyzed in this study (Fig. S1D). The four models were fitted using the phylogeny proper to each gene and the phylogeny of all concatenated loci.

Because individual locus analyses could lack statistical power, we performed additional analyses using two alternative procedures: summing log-likelihoods of individual loci and analyzing concatenated loci. In the former approach, loci were analyzed separately using their own phylogeny (Table S2). Phylogenies and substitution patterns can therefore vary among loci, which is suitable. However, this approach can easily attribute statistical significance if log-likelihoods are directly summed. To correct for this, sums were conditional to the mating-system pattern, that is, when ω ratios were greater in selfing than outcrossing species we summed the log-likelihood of model (4); otherwise we summed the log-likelihood of model (3). Correspondingly, when GC* ratios were greater in outcrossing than selfing species, we summed the log-likelihood of model (4); otherwise we summed

the log-likelihood of model (3). In the latter approach, concatenated loci were analyzed using the tree presented in Figure 2. This approach constrains analyses to one phylogeny and substitution patterns are assumed to be the same along the concatenate, which is certainly disadvantageous. However, it is more robust than the former approach.

For the analyses of concatenated loci, loci were first concatenated using a homemade Perl script. When data for a given species and locus were missing, the alignment contained an *N*-filled sequence as long as the size of the locus. Seven loci were excluded from concatenates because they lacked several species and inserted large *N*-filled fragments in the final alignment. These loci were LOC_Os01g21160, LOC_Os01g53720, LOC_Os01g60230, LOC_Os01g61720, LOC_Os01g68770, *PinA* and *PinB*. The final alignment of the 20 concatenated loci contained 18,729 bp. In addition to the four hierarchical models described above, we fitted four extra models using concatenates (numbering consecutive after the models described above). (5) A model in which outcrossing was assumed for all internal branches within Triticeae, excepting the selfing branch linking *H. bogdanii* and *H. marinum* (Fig. S1E). (6) A model in which a different ratio was estimated for actual outcrossing, actual selfing, and species with unknown, potentially mixed-mating systems (Fig. S1F). (7) The one-ratio-per-clade model estimated one substitution ratio per clade, one ratio for the internal branches leading to the different clades, and one ratio for the outgroup (Fig. S1G). (8) The mating-system-and-clade model was similar to model (7) but estimated three ratios within each clade: selfing, outcrossing, and internal branches of each clade; this model estimated 15 parameters instead of 17 because clade I (*Psathyrostachys*) had only one ratio to estimate (Fig. S1H).

Log-likelihood sums and analyses of concatenated loci were performed across all loci ($N = 27$) and, additionally, across five different groups of loci. Loci were pooled according to relevant genomic parameters, as follows. First, they were grouped according to their total GC content. Loci for which the GC content was lower or equal than the median were considered as GC-poor ($N = 14$); otherwise as GC-rich ($N = 13$). Second, to test for the impact of recombination on substitution rates, we pooled loci according to their physical location on chromosomes. For the chromosome 3, loci were considered as centromeric if located at a relative distance to the centromere lower than 0.70 ($N = 12$); otherwise as telomeric (note that *GSP*, which is a telomeric locus located on chromosome 5, was included in the analysis of telomeric loci, but *PinA* and *PinB* were excluded because they introduced large *N*-filled fragments; $N = 12$) (Table 1). Finally, we pooled loci congruent with the phylogeny of all concatenated loci ($N = 17$) to verify that results are not due to specific incongruent loci.

We tested for correlations among the relative distance to the centromere, ω, and GC*. Nonparametric correlations were tested using estimates of substitution rates for Triticeae branches, obtained from the two-ratio model. Likewise, correlations were tested for outcrossing and selfing branches, obtained from the four-ratio model. Analyses included ω and its component terms (d_N and d_S), GC_{123}* and GC_3*. The α-level was adjusted for multiple comparisons (false discovery rate control; Verhoeven et al. 2005). Correlation analyses were performed with JMP 3.2.1 (SAS Institute).

Results

EVOLUTIONARY TRANSITIONS OF THE MATING SYSTEM IN TRITICEAE

It has been shown that the distribution of selfing rates in wind-pollinated plants is basically bimodal (Schemske and Lande 1985; Goodwillie et al. 2005). Therefore, a dichotomous classification of the mating system seems, a priori, appropriate in Triticeae. Accordingly, we were able to distinguish nine selfing and 10 outcrossing species in our dataset (Fig. 1; Table S3). We defined selfing species as those exhibiting large autonomous seed set and low pollen/ovule ratio (or small anthers), whereas outcrossing species were those exhibiting the inverse pattern. Among outcrossers, we distinguished annual ($N = 5$) and perennial ($N = 5$) species, whereas all selfing species were annual. Although this classification simplifies analyses, it is clear from Figure 1 that four species (*E. bonaepartis*, *H. bogdanii*, *H. piliferum*, and *T. caput-medusae*), classified as selfers, had intermediate autonomous seed sets (0.2–0.5). These species could represent mixed maters rather than selfing species. Alternatively, it could be that glass-house conditions were not optimal for pollination in these species. These uncertainties are taken into account in analyses (see below).

A parsimony reconstruction of ancestral states, using the dichotomous classification of the mating system and the multigenic phylogeny of Triticeae, indicates that there were six transitions of the mating system in the history of Triticeae: five independent transitions from outcrossing to selfing, and one transition from selfing to outcrossing. According to this analysis, the most parsimonious state of the mating system of the ancestor of Triticeae is outcrossing (Fig. 2A). Consistently, both ML and MCMC analyses suggest that the likely ancestral state of the majority of the nodes in the phylogeny is outcrossing, and transitions would mainly proceed from outcrossing to selfing (Fig. 2B). Both analyses suggest that the log-likelihood of the unrestricted model, allowing transitions from outcrossing to selfing (q_{os}) and from selfing to outcrossing (q_{so}), was not significantly better than that of the restricted model allowing only transitions from outcrossing to selfing ($q_{so} = 0$). However, the log-likelihood was significantly better than that of the alternative restricted model in which $q_{os} = 0$ (Table 2). Analyses considering *E. bonaepartis*,

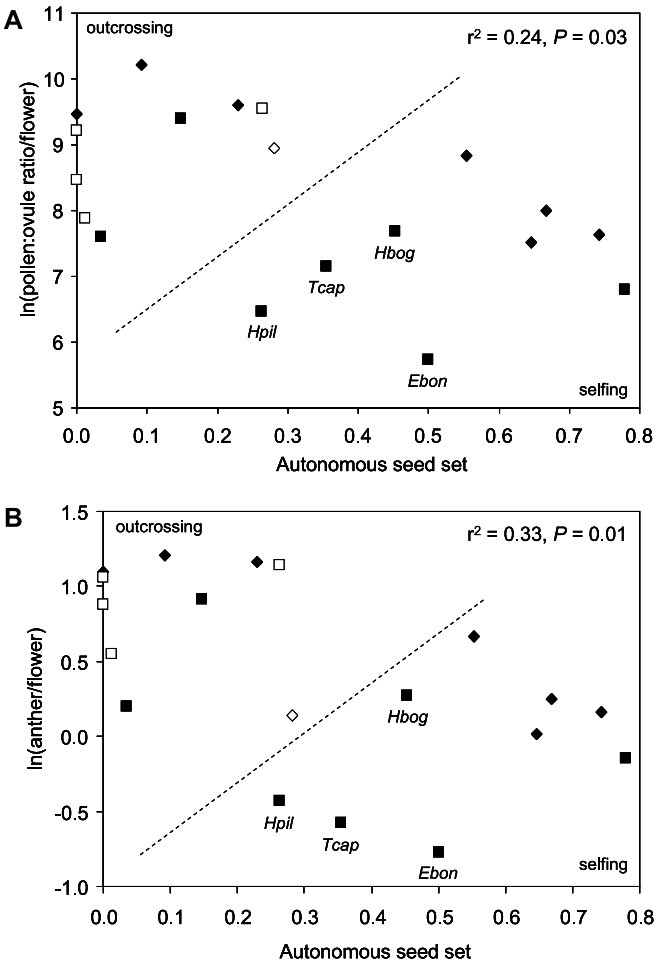

Figure 1. Correlations between the autonomous seed set and the log-pollen/ovule ratio (A) and the anther size (B). Note that the pollen/ovule ratio and anther size were scaled to the flower size. Filled symbols, annual species; open symbols, perennial species; diamonds, species with known mating system; squares, species with unknown mating system. Dashed lines are given for illustrative purposes to differentiate outcrossing and selfing species in our dataset. Ebon, *Eremopyrum bonaepartis*; Hbog, *Hordeum bogdanii*; Hpil, *Heteranthelium piliferum*; Tcap, *Taeniatherum caput-medusae*.

H. bogdanii, *H. piliferum*, and *T. caput-medusae* as undetermined mating systems show basically the same results: the best model describing transitions in the mating system across the Triticeae phylogeny is that in which selfing-to-outcrossing transitions are neglected (Table 2).

IMPACT OF THE MATING SYSTEM AND RECOMBINATION ON RATES OF PROTEIN EVOLUTION

A total of 23,574 bp of coding regions have been sequenced, aligned, and analyzed. Results on base composition, substitution rates, and other relevant parameters for all sequenced loci are presented in Table 1. ω ratios were estimated using the phylogeny proper of each gene. Mean ω ratios for outcrossing and selfing branches of the phylogeny are 0.15 and 0.14, respectively (ranges: 0.0001–0.53 and 0.0001–0.38, respectively). Selfing species exhibit ω ratios greater than outcrossing species at 13 loci, consistent with theoretical predictions linked to their expected lower effective population size and lower recombination rate. However, 13 other loci show the inverse pattern and one locus (*PSY2*) exhibits the same ω ratio in the two mating systems (Fig. 3A). In most cases, differences in ω ratios between the two states of the mating system are not significant. The difference between selfers and outcrossers is not significant either when log likelihoods are combined (Table 3). One gene (LOC_Os01g48720) shows evidence of positive selection (results not shown). However, the results are unchanged when this gene is removed.

Because lack of significance could be due to a limited statistical power of individual locus analyses, we also performed analyses on sequences concatenated in several ways (see Materials and Methods). ω ratios for concatenated sequences were estimated using the phylogeny of all concatenated loci (Fig. 2). Contrary to the theoretical expectations, ω ratios are always greater in outcrossing than selfing branches, although differences are not significant (Table 3). Effective population sizes and/or selective constraints may have varied along the phylogeny for other reasons than a shift in the mating system. To control for the potential impact of the phylogenetic history on protein evolution, we additionally analyzed substitution rates per clade (Fig. 2) using concatenated sequences (models 7 and 8; see Materials and Methods, and Figs. S1G and H). Although this analysis does not reveal any statistically significant difference between selfing and outcrossing, it gives a more detailed idea about variation in substitution rates across the phylogeny. In most clades and the majority of concatenated loci, ω ratios are greater in outcrossing than selfing branches (Table 4).

Because some species have intermediate phenotypic traits between selfing and outcrossing species (Fig. 1), we performed additional analyses in which ω ratios were estimated for selfing, outcrossing, and species with unknown mating system (Table 5). In all cases but GC-rich loci, ω ratios are greater in outcrossing than selfing species, although statistical significance was not reached. The robustness of our analyses was also tested by (1) differentiating perennial and annual outcrossers, (2) performing single-locus analyses with the phylogeny of all concatenated sequences instead of the gene phylogeny, and (3) assuming outcrossing as the mating system in internal nodes. Results are mainly unchanged (Table S4). To summarize, our results on protein evolution, estimated by the ω ratios, provide no support to the hypothesis that selfing species have accumulated more deleterious mutations than outcrossing ones (second assumption of the dead-end hypothesis of selfing evolution).

However, we detect the expected effect of recombination on protein evolution predicted by the same body of theory.

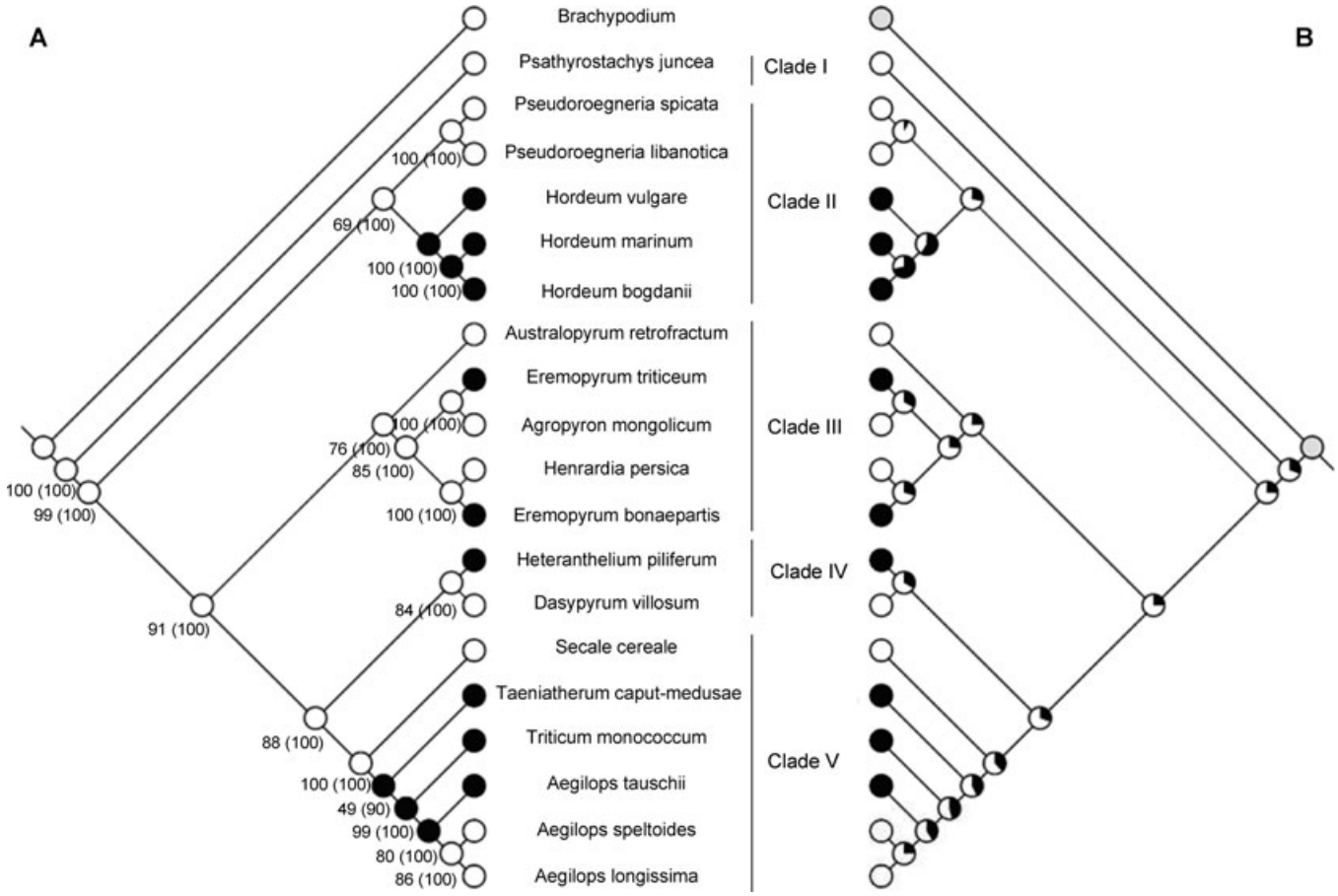

Figure 2. Evolutionary transitions of the mating system in Triticeae. (A) Parsimony reconstruction. (B) Bayesian reconstruction. Trees in (A) and (B) correspond to the multigenic maximum-likelihood tree. Values under the nodes in panel (A) represent bootstrap and Bayesian posterior supports (between parentheses). Open circles, outcrossing; filled circles, selfing; gray circles, undetermined.

Recombination gradients are strong in wheat (Akhunov et al. 2003; Saintenac et al. 2009), as well as in other Triticeae species (Dubcovsky et al. 1996; Luo et al. 2000, 2005). Most crossovers are physically located in the distal one-third of chromosome arms and recombination increases exponentially from the centromere to telomeres (Akhunov et al. 2003). We obtained 21 loci, of the 27 studied loci, from chromosome 3 known to be collinear among wheat, *Aegilops* (Zhang et al. 2001), *B. dystachion* (Huo et al. 2006), and rice (Sorrells et al. 2003; Munkvold et al. 2004; Haudry et al. 2008). We used the relative distance to the centromere of these 21 loci to assess the impact of recombination intensity in substitution rates. For these loci, ω ratios are negatively correlated with the relative distance to the centromere (Spearman's rho $= -0.47$, $P = 0.03$ if including the locus under positive selection; rho $= -0.44$, $P = 0.05$ if excluding it; Fig. 4A). The same trend illustrating the recombination impact on protein evolution is observed when estimating ω ratios in selfing and outcrossing branches separately (outcrossing: rho $= -0.38$, $P = 0.10$; selfing: rho $= -0.49$, $P = 0.03$; analyses excluding the locus under positive selection). Consistent with this, ω ratios are greater for centromeric than telomeric loci (Table 3). Instead of using the phylogeny of each gene, we redid the analyses by imposing the phylogeny of all concatenated sequences. As above, ω ratios are negatively correlated with the relative distance to the centromere (ω Triticeae: rho $= -0.46$, $P = 0.04$; ω outcrossing: rho $= -0.40$, $P = 0.08$; ω selfing: rho $= -0.43$, $P = 0.06$; analyses excluding the locus under positive selection), demonstrating that correlations are robust and do not depend on the phylogenetic framework. In the light of the results on recombination, the lack of evidence of a mating system impact on protein evolution is probably not due to a weak statistical power of our dataset.

IMPACT OF MATING SYSTEM AND RECOMBINATION ON GC COMPOSITION

gBGC is highly sensitive to the effective rate of recombination (in heterozygous state), hence to the level of outcrossing (Marais 2003; Marais et al. 2004). It can interfere with selection and fix slightly deleterious GC alleles (Galtier and Duret 2007; Galtier et al. 2009). This can lead to a spurious increase in the ω ratio in highly recombining regions and genomes. In grasses, gBGC is

Table 2. **Models testing transitions between mating systems across the Triticeae phylogeny. Two sets of results are shown: one assuming a dichotomous classification of the mating system (species are either outcrossing or selfing) and the other in which the mating system of *E. bonaepartis*, *H. bogdanii*, *H. piliferum*, and *T. caput-medusae* is unknown. ML, maximum likelihood; MCMC, Bayesian Markov Chain Monte Carlo; lnL, log-likelihood (mean log-likelihoods for ML analyses, and harmonic mean of log-likelihoods for MCMC analyses); Dev, deviance (−2×log-likelihood); *P*, *P*-value; q_{os}, probability of outcrossing-to-selfing transitions; q_{so}, probability of selfing-to-outcrossing transitions; q_o, probability of outcrossing at the root of the tree; q_s, probability of selfing at the root of the tree.**

Model	lnL	Dev (*P*)	q_{os}	q_{so}	q_o	q_s
ML analyses (dichotomous classification of the mating system)						
ML unrestricted	−12.36		90.63	81.63	0.71	0.29
ML q_{so}=0	−12.54	0.36 (0.54)	42.67	0.00	1.00	0.00
ML q_{os}=0	−15.84	6.95 (<0.01)	0.00	44.65	0.00	1.00
ML analyses (four species with unknown mating system)						
ML unrestricted	−8.90		77.02	130.92	0.71	0.29
ML q_{so}=0	−9.19	0.58 (0.45)	27.51	0.00	1.00	0.00
ML q_{os}=0	−12.19	6.58 (0.01)	0.00	62.34	0.00	1.00
MCMC analyses (dichotomous classification of the mating system)						
MCMC unrestricted	−13.06		65.64	58.38	0.72	0.28
MCMC q_{so}=0	−13.32	0.51 (0.47)	49.48	0.00	1.00	0.00
MCMC q_{os}=0	−16.57	7.02 (<0.01)	0.00	50.79	0.00	1.00
MCMC analyses (four species with unknown mating system)						
MCMC unrestricted	−9.71		52.26	62.86	0.73	0.26
MCMC q_{so}=0	−10.29	1.15 (0.28)	35.73	0.00	1.00	0.00
MCMC q_{os}=0	−12.77	6.12 (0.01)	0.00	65.12	0.00	1.00

supposed to occur (Glémin et al. 2006; Haudry et al. 2008). We thus investigated if GC content evolution is affected by shifts in the mating system to test whether gBGC could explain why we are unable to detect differences in protein evolution between the two mating systems. Indeed, gBGC could increase the ω ratio of outcrossing species as previously proposed for *S. cereale* and *Ae. speltoides* (Haudry et al. 2008).

The average GC content at all codon positions (GC_{123}) across the 27 sequenced loci is 0.47 (median: 0.46; range: 0.41–0.57). Average GC_1, GC_2, and GC_3 are, respectively, 0.54, 0.42, and 0.46 across all loci, respectively (Table 1). Using a maximum likelihood framework (Dutheil and Boussau 2008) and the phylogeny proper of each gene (Table S2), we tested if outcrossing branches experienced more AT → GC substitutions than selfing ones, by estimating the equilibrium GC content (GC*). Thirteen loci (out of 27) exhibit GC_{123}* was greater in outcrossing than selfing branches, consistent with theoretical predictions (one significant). However, the 14 other loci exhibit the inverse pattern (six significant) (Fig. 3B). Analysis of GC_3* shows very similar results: 10 loci exhibit GC_3* greater for outcrossing than selfing branches (one significant), whereas all other loci exhibit the opposite pattern (four significant).

Like for protein evolution, we performed additional analyses of substitutions toward G or C bases to avoid potentially limited statistical power. The phylogeny of all concatenated loci (Fig. 2) was used to estimate GC* in concatenated sequences. Neither log-likelihood sums across loci nor concatenated sequences reveal significant differences between selfing and outcrossing lineages for their GC-content evolution. This is the case for GC_{123} and GC_3 analyses. Excepting the concatenate of telomeric loci, we find that GC_{123}* and GC_3* are greater in selfers than in outcrossers, although mostly not significant.

Analyses of concatenated sequences per clade reveal important variations in GC_{123}* and GC_3* across the phylogeny. In all cases, the model including both clade and mating system effects (model 8; see Materials and Methods and Fig. S1H) is significantly better than the model including only a clade effect (model 7; see Materials and Methods and Fig. S1G) (Table 4). Although each clade has a different baseline substitution rate toward G or C, there is no general trend for outcrossing branches being associated with higher GC*. Two extreme cases are illustrated by clades II (*Hordeum* and *Pseudoroegneria*) and V (*Secale*, *Taeniatherum*, *Triticum* and *Aegilops*). In the former, GC_{123}* and GC_3* are always greater in selfing than outcrossing branches, whereas in the latter GC_{123}* and GC_3* are greater for outcrossing than selfing branches, except for GC_3* of GC-rich and congruent loci (Table 4). As for protein-evolution analyses, we tested the robustness of our analyses in several ways. Although GC_{123}* were greater in outcrossing than selfing branches when assuming the former as the mating system of internal branches, the difference was not statistically significant. In all other cases, results were mainly unchanged (Table S4).

Taken together, we have no evidence suggesting enrichment in G and C alleles in the genomes of outcrossing compared to

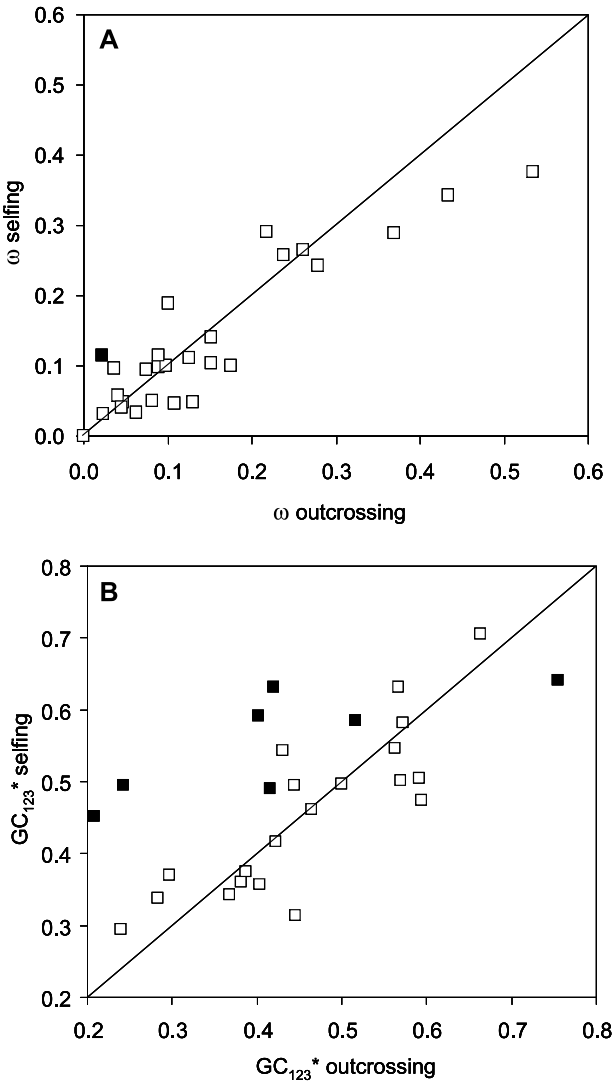

Figure 3. Substitution patterns of individual loci for selfing and outcrossing branches of the phylogeny. (A) ω ratios. (B) GC_{123}* ratios. The diagonal line represents equal substitution rates between selfing and outcrossing. Filled squares, statistically significant differences between mating systems; open squares, no statistically significant differences between mating systems. Significance was determined through likelihood-ratio tests.

selfing species in Triticeae. As for protein evolution, this result seems not due to a limited statistical power. In agreement with either the gBGC hypothesis or selection favoring GC codons, we detect a positive and significant correlation between GC_{123}* and the relative distance to the centromere (Triticeae branches: rho = 0.47, $P = 0.04$; analysis excluding the locus under positive selection; Fig. 4B). The same pattern is found when analyzing outcrossing and selfing branches separately (outcrossing: rho = 0.53, $P = 0.02$; selfing: rho = 0.38, $P = 0.10$; analyses excluding the locus under positive selection), and GC_3* instead of GC_{123}*, although in this case significance disappeared (rho = 0.31, $P = 0.18$; analysis excluding the locus under positive selection). Accordingly, GC_{123}* and GC_3* are greater in telomeric than centromeric loci (Table 3). Finally, contrary to the "genomic Achilles' heel" hypothesis, GC* and ω ratios are negatively correlated at the Triticeae scale (GC_{123}*: rho = −0.51, $P = 0.008$; GC_3*: rho = −0.47, $P = 0.01$; analyses excluding the locus under positive selection). This is likely due to the fact that both parameters are correlated to recombination. Therefore, a higher ω ratio in outcrossing branches seems not due to higher substitutions toward GC.

Discussion

Using a phylogenetic framework, we combined morphological characters linked to the mating system and molecular evolution analyses to test for the two assumptions on which the theory that selfing is an evolutionary dead end is based on. First, we assessed the validity of irreversibility of transitions from outcrossing to selfing. Second, we tested for the hypothesis that selfing lineages should accumulate slightly deleterious mutations, considered as a prelude to their extinction. Altogether, our results provide insight into the tempo and mode of evolution of self-fertilization in hermaphroditic grasses.

ARE MATING-SYSTEM TRANSITIONS IRREVERSIBLE IN TRITICEAE?

The first assumption of the dead-end hypothesis of selfing evolution is that selfing lineages cannot revert to outcrossing. Parsimony and probabilistic models suggest that outcrossing is the likely ancestral state in Triticeae and indicate that transitions have mainly occurred from outcrossing to selfing. Given the current phylogeny of the tribe, most internal nodes have high probability of being outcrossing (0.54–0.87) and only nodes within *Hordeum* (clade II) have higher probabilities of being selfing than outcrossing (Fig. 2B). Although parsimony reconstruction indicates one point in which the reverse transition could have occurred (the branch leading to *Ae. speltoides* and *Ae. longissima*; Fig. 2A), both ML and MCMC methods suggest that the evolution of the mating system in Triticeae is best described by a model in which selfing-to-outcrossing transitions are neglected. Our results agree with the prediction that selfing species evolve from outcrossing ancestors, although they do not reject the reverse transition. Similar results have been found in most previous studies performed in plants (Armbruster 1993; Barrett et al. 1996a; Kohn et al. 1996; Goodwillie 1997; Schoen et al. 1997; Igic et al. 2006; but see Bena et al. 1998; Ferrer and Good-Avila 2006).

Although our results are consistent with the first assumption of the evolutionary dead-end hypothesis, they must be taken with caution. First, methods for reconstructing ancestral states assume that the phylogenetic tree reflects the true topology of the tribe (Pagel 1999). Triticeae is a tribe in which phylogenetic reconstruction has been particularly difficult (Kellogg et al. 1996).

Table 3. Effect of the mating system on substitution parameters. Dev, deviance; *P*, *P*-value; Out, outcrossing; Self, selfing. Congruent genes are those showing <5% of incongruence in Table 1.

	ω			GC_{123}*			GC_3*		
	Dev (*P*)	Out	Self	Dev (*P*)	Out	Self	Dev (*P*)	Out	Self
Sum of log-likelihoods									
All loci	14.8 (0.96)	0.148	0.136	20.4 (0.77)	0.450	0.481	22.2 (0.68)	0.436	0.472
GC-poor	5.4 (0.97)	0.155	0.142	6.6 (0.92)	0.415	0.460	4.9 (0.98)	0.392	0.421
GC-rich	9.4 (0.67)	0.152	0.140	13.7 (0.32)	0.481	0.501	17.3 (0.14)	0.477	0.524
Centromere	6.5 (0.84)	0.154	0.143	8.1 (0.70)	0.395	0.450	4.9 (0.94)	0.381	0.424
Telomeres	8.1 (0.70)	0.157	0.148	12.2 (0.35)	0.505	0.497	17.3 (0.10)	0.480	0.478
Congruent genes	13.4 (1.00)	0.127	0.131	13.6 (0.80)	0.437	0.495	16.8 (0.27)	0.427	0.484
Concatenated sequences									
All loci	1.7 (0.19)	0.132	0.118	1.4 (0.24)	0.452	0.466	5.4 (0.08)	0.427	0.443
GC-poor	1.1 (0.29)	0.143	0.128	0.8 (0.37)	0.417	0.431	0.0 (1.00)	0.389	0.389
GC-rich	0.4 (0.54)	0.111	0.102	11.0 (<0.001)	0.512	0.518	1.2 (0.27)	0.501	0.538
Centromere	1.3 (0.26)	0.166	0.147	2.2 (0.14)	0.425	0.452	1.4 (0.24)	0.381	0.409
Telomeres	0.1 (0.79)	0.122	0.118	0.0 (1.00)	0.516	0.507	0.0 (1.00)	0.503	0.501
Congruent genes	0.4 (0.81)	0.120	0.118	8.4 (0.004)	0.462	0.499	5.4 (0.02)	0.438	0.474

However, the multigenic phylogeny we obtained is the most robust phylogeny to date in this tribe in terms of the number of sequenced genes (27 compared to one to three genes in previous studies). We are confident that the phylogenetic relationships among genera reflect the predominant phylogenetic signal as most nodes have high bootstrap and posterior Bayesian support. Second, the ability to test for irreversibility depends on the size of the phylogeny, on when the trait is first gained, and on the overall rates of gain and loss of the trait (Sanderson 1993; Ferrer and Good-Avila 2006). It is possible that including more species in the current phylogeny

Table 4. Effect of the mating system on substitution ratios per clade. Clades are depicted in Figure 2. Dev, deviance; *P*, *P*-value; Out, outcrossing; Self, selfing. Congruent genes are those showing <5% of incongruence (Table 1).

Concatenate	Dev (*P*)	Clade I	Clade II		Clade III		Clade IV		Clade V	
		Out	Out	Self	Out	Self	Out	Self	Out	Self
ω										
All loci	12.1 (0.14)	0.129	0.139	0.109	0.131	0.124	0.102	0.177	0.136	0.119
GC-poor	14.5 (0.07)	0.152	0.155	0.114	0.134	0.128	0.107	0.164	0.147	0.137
GC-rich	6.3 (0.62)	0.094	0.114	0.099	0.121	0.113	0.076	0.215	0.117	0.095
Centromere	11.1 (0.20)	0.115	0.194	0.136	0.182	0.176	0.095	0.171	0.183	0.141
Telomeres	3.3 (0.91)	0.136	0.097	0.095	0.111	0.123	0.155	0.221	0.126	0.129
Congruent genes	14.7 (0.06)	0.096	0.086	0.105	0.137	0.145	0.042	0.171	0.138	0.113
GC_{123}*										
All loci	28.6 (<0.001)	0.579	0.398	0.496	0.450	0.446	0.425	0.483	0.467	0.436
GC-poor	24.0 (<0.01)	0.549	0.367	0.459	0.398	0.423	0.428	0.490	0.409	0.380
GC-rich	20.4 (<0.01)	0.593	0.465	0.554	0.530	0.491	0.434	0.326	0.550	0.518
Centromere	18.6 (0.02)	0.509	0.389	0.492	0.448	0.426	0.378	0.477	0.412	0.402
Telomeres	25.0 (<0.01)	0.610	0.456	0.533	0.454	0.509	0.534	0.474	0.564	0.495
Congruent genes	41.4 (<0.001)	0.577	0.403	0.544	0.464	0.493	0.496	0.609	0.506	0.470
GC_3*										
All loci	24.4 (<0.01)	0.467	0.342	0.455	0.434	0.416	0.393	0.465	0.454	0.425
GC-poor	20.8 (<0.01)	0.490	0.301	0.393	0.375	0.377	0.395	0.468	0.423	0.366
GC-rich	44.0 (<0.0001)	0.553	0.452	0.562	0.543	0.500	0.410	0.348	0.506	0.518
Centromere	27.8 (<0.001)	0.415	0.291	0.435	0.437	0.372	0.299	0.480	0.386	0.368
Telomeres	20.0 (0.01)	0.692	0.436	0.519	0.436	0.491	0.542	0.341	0.522	0.506
Congruent genes	44.8 (<0.0001)	0.546	0.356	0.499	0.448	0.439	0.369	0.588	0.423	0.436

Table 5. Effect of the mating system on substitution parameters. In this analysis, different substitution rates are estimated for actual selfing (Self), outcrossing (Out), and species with unknown mating system (Unknown). Dev, deviance; *P*, *P*-value; Out, outcrossing; Self, selfing. Congruent genes are those showing <5% of incongruence (Table 1).

	ω				GC_{123}*				GC_3*			
	Dev (*P*)	Out	Self	Unknown	Dev (*P*)	Out	Self	Unknown	Dev (*P*)	Out	Self	Unknown
All loci	0.50 (0.48)	0.131	0.120	0.127	0.2 (0.65)	0.465	0.465	0.446	0.6 (0.44)	0.437	0.435	0.426
GC-poor	0.86 (0.35)	0.154	0.131	0.132	0.0 (1.00)	0.463	0.419	0.422	1.8 (0.18)	0.429	0.381	0.375
GC-rich	0.26 (0.61)	0.095	0.103	0.115	1.2 (0.27)	0.495	0.540	0.497	0.0 (1.00)	0.478	0.531	0.518
Centromere	0.14 (0.70)	0.129	0.121	0.108	6.0 (0.01)	0.416	0.458	0.434	1.4 (0.24)	0.367	0.411	0.394
Telomeres	0.21 (0.65)	0.130	0.120	0.108	2.2 (0.14)	0.567	0.523	0.478	10.0 (0.002)	0.554	0.508	0.462
Congruent genes	1.54 (0.21)	0.122	0.114	0.123	3.0 (0.08)	0.462	0.508	0.483	0.6 (0.44)	0.442	0.467	0.449

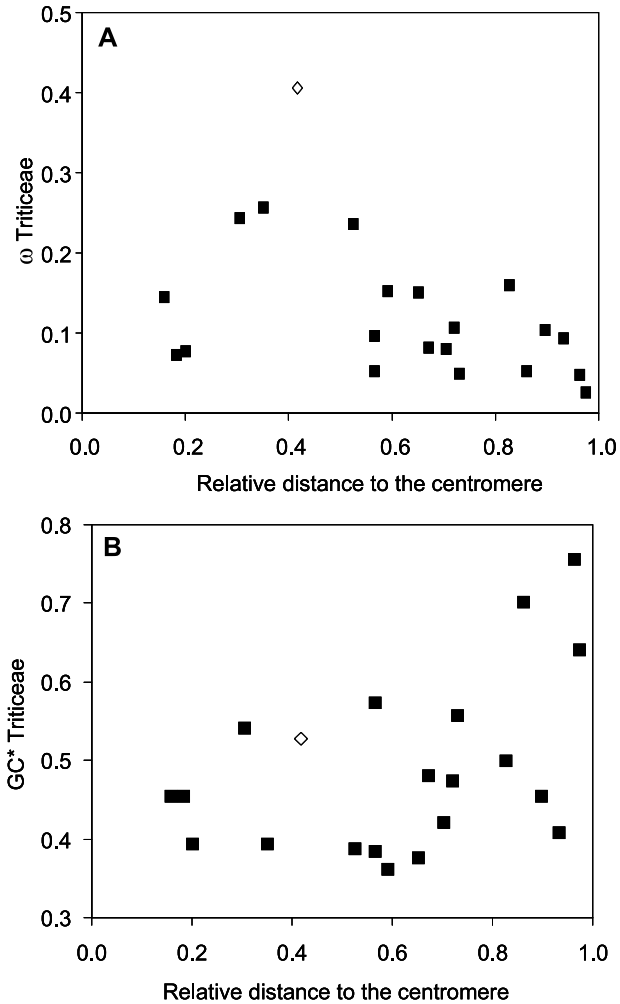

Figure 4. Correlations between the relative distance to the centromere and substitution rates for loci located on chromosome 3. (A) ω per locus. (B) GC_{123}* per locus. Substitution rates are calculated from the two-ratio model (see Materials and Methods). The gene showing evidence of positive selection is depicted (open diamond) but was not taken into account in correlation analyses.

alters the picture we provide about transitions in the mating systems. Finally, Igic et al. (2006) and Goldberg and Igic (2008) showed that using only the character states of extant species to infer the ancestral states can lead to spurious results. However, it tends to overestimate reversible transition, such that the conclusion of irreversible evolution from outcrossing to selfing is rather robust.

Divergence of the ancestor of Triticeae is estimated to have occurred ~12–15 million years ago (Mya), given the established idea that wheat (*Triticum*)-barley (*Hordeum*) divergence have occurred ~10 Mya (Dvorak and Akhunov 2005). According to our analyses, in this relatively short-time period, Triticeae experienced several independent transitions from outcrossing to selfing, suggesting that mating systems are evolutionary labile features in this tribe. As several independent outcrossing-to-selfing

transitions can be sampled across the phylogeny, Triticeae seems an appropriate group to evaluate the impact of the mating system on genome evolution.

DO SELFERS ACCUMULATE MORE DELETERIOUS MUTATIONS THAN OUTCROSSERS?

The second assumption of the dead-end hypothesis of selfing evolution is that selfing species would become extinct because of the accumulation of slightly deleterious mutations and/or limited potential for adaptation. The underlying condition is that the effective population size and effective recombination rate are reduced in selfers. We assess this through the analysis of substitution patterns on protein-coding sequences between selfing and outcrossing species: we expect nonsynonymous substitutions (measured through ω ratios) to accumulate more in selfers than outcrossers.

We find that ω ratios were negatively correlated with the relative distance to the centromere and greater in centromeric than telomeric regions. Similar patterns showing the impact of recombination on selection efficacy have been previously shown in *Drosophila* (Presgraves 2005; Haddrill et al. 2007; Betancourt et al. 2009). Because recombination gradients are strong in Triticeae, our results suggest that selection is relaxed when recombination is low. Because of their reduced effective recombination rate, selection is also expected to be relaxed in lowly recombining genomes, especially in self-fertilizing species. However, none of our analyses reveal any clear difference between the two mating systems for the efficacy of selection on protein-coding sequences. On the contrary, most genes show greater ω ratios for outcrossing than selfing branches, and the analysis per clade only reveal a tendency for increased ω ratios in selfing relative to outcrossing in clade IV (*Dasypyrum–Heteranthelium*), although the difference is not statistically significant. The dichotomous classification of the mating system could be responsible for these results if mixed-mating species have been pooled with actual selfing species (Fig. 1). However, very similar results are obtained when a different ratio is estimated for species with unknown mating system and species with known mating systems are compared (Table 5). It seems therefore that results are not much affected by our mating-system classification based on phenotypic correlates.

Similar results have been previously obtained in other angiosperms and in nematodes: no difference in substitution patterns between selfing and outcrossing species was observed in *Arabidopsis* (Wright et al. 2002), a subset of four Triticeae species (*Ae. speltoides*, *S. cereale*, *T. monococcum* and *T. urartu*; Haudry et al. 2008) and *Caenorhabditis* (Cutter et al. 2008). In a thorough analysis of several families, Glémin et al. (2006) found only a weak tendency for selection to be less effective in selfers at removing weakly deleterious alleles. Altogether, current studies do not support the hypothesis that selfing lineages accumulate more deleterious mutations than outcrossing ones. A possible explanation for this could be that selfing is not a fixed strategy in natural populations. It is known that well-characterized selfing species, such as *Arabidopsis thaliana* and *Caenorhabditis elegans*, do not self-fertilize at 100% (Savolainen et al. 2000; Barrière and Félix 2005; Morran et al. 2009). More or less long-term episodes of excess cross- or self-fertilization might affect to an unknown extent the expected substitution patterns of molecular evolution. However, the effective population size is expected to be reduced even in partially self-fertilizing populations if background selection is strong (Charlesworth et al. 1993; Nordborg et al. 1996), or because of limited gene flow and extinction–recolonization dynamics (Ingvarsson 2002). Reduction in the effective population size in selfers is also a general trend observed in polymorphism data (Hamrick and Godt 1996; Nybom 2004; Glémin et al. 2006; Foxe et al. 2009; Guo et al. 2009).

More recently, following Galtier and Duret (2007), Haudry et al. (2008) suggested that gBGC could explain higher ω ratios in outcrossers than in selfers. Contrary to the dead-end hypothesis, outcrossers, and not selfers, would suffer from a gBGC induced load (the "genomic Achilles' heel"). However, analyses of GC-content dynamics gave a picture similar to the one given by protein evolution: recombination but not mating system affects molecular evolution in Triticeae. The pattern does not depend on our dichotomous classification of the mating system (compare results for selfing and outcrossing between Tables 3 and 5). Only one clade (clade V, i.e., *Secale*, *Taeniatherum*, *Triticum* and *Aegilops*) shows a consistent pattern of higher GC^* in outcrossing than selfing branches, in agreement with previous results obtained in few Triticeae species all belonging to this clade (Haudry et al. 2008). The significant negative correlation between ω and GC^* does not support the Achilles' heel hypothesis at this scale. The lack of effect of mating system on GC-content dynamics is even more surprising because gBGC should be more strongly affected by mating system than selection (Haudry et al. 2008). Selection is mainly affected by variation in the effective population size associated with shifting in the mating system. gBGC, on the other hand, is expected to vary much more dramatically because it vanishes under homozygosity, hence selfing (Marais et al. 2004; Haudry et al. 2008). Under gBGC, GC-content and GC^* should be good predictors of the mating system (Glémin et al. 2006; Haudry et al. 2008).

CONSERVED RECOMBINATION PATTERNS VERSUS RAPID AND RECENT SHIFTS TOWARD SELFING

Taken together, our results strongly support the view that shifts in the mating system are rapid and that selfing is of recent origin in Triticeae, whereas rough recombination patterns (e.g., centromere vs. telomeres) are conserved at the scale of Triticeae. This is not surprising given the strong collinearity of genes among wheat—*Aegilops*, *Brachypodium*—and rice. Hence, it seems reasonable

to think that rough recombination patterns have rested mostly unchanged across the evolution of Triticeae, allowing to detect differences in ω and GC^* ratios between regions of low and high recombination.

Within the Triticeae, selfing is probably too recent in many terminal branches to detect sufficient differences with outcrossing lineages. In our analyses, ω and GC^* are averages over the past history of mating systems along branches. If selfing is recent, ω and GC^* could mainly reflect the substitution history of an outcrossing lineage while we assigned a selfing status to branches leading to extant selfing species. Studies in the Brassicaceae have shown that selfing may be of very recent origin, as in *A. thaliana* ($\sim$1 Mya; Tang et al. 2007) and *Capsella rubella* ($\sim$25,000 years ago; Foxe et al. 2009; Guo et al. 2009). In Triticeae, there is evidence suggesting that *Hordeum bulbosum*, a self-incompatible species not analyzed in the present study, branches out in one of the first divergent subclades of the genus, close to *H. vulgare* (Blattner 2004). It could mean that the branch leading to *Hordeum* has spent most of the time since the common ancestor with an outcrossing breeding system, while we hypothesized in our analyses that the whole branch has been evolving under selfing. Consistent with this, GC^* ratios were higher in outcrossing than selfing branches when declaring outcrossing as the mating system of internal branches (Table S4). If selfing recently evolved, we may have missed other transitions. Accordingly, the difference between our study and Haudry et al. (2008), where the effect of mating system on GC^* was strong, can be explained because they studied a couple of self-fertilizing sister species, such that they could reasonably assume that selfing persisted from the common ancestor to the extant species.

CONCLUSION: IS SELFING AN EVOLUTIONARY DEAD END IN TRITICEAE?

Our results suggest that extant selfing species have mostly evolved from outcrossing ancestors in Triticeae. Even if reversibility in mating-system transitions could not be rejected, our data are not in contradiction with the first assumption of the evolutionary dead-end hypothesis. On the other hand, although we do not provide direct support to the prediction that selfing species accumulate more deleterious mutations than outcrossing ones, our results are consistent with selfing lineages becoming extinct faster than outcrossing ones. The recent origin of self-fertilization has already been suggested in other groups of plants (Barrett et al. 1996b; Kohn et al. 1996; Schoen et al. 1997; Bena et al. 1998; Wright et al. 2002; Foxe et al. 2009; Guo et al. 2009) and animals (Kiontke et al. 2004; Cutter et al. 2008). Here, thanks to the contrasted effects of recombination and current mating systems, we can calibrate the tempo of genome evolution. We thus have sound arguments to state that self-fertilization may be of recent origin in Triticeae and inbreeding may have persisted for insufficient time for large differences between the two states of the mating system to evolve. To evaluate the second prediction of the dead-end hypothesis, one needs to determine when the shift in the mating system has taken place along a particular branch of the phylogeny. Ideally, one would like to compare polymorphism and divergence data in groups of several sister species with varying mating systems.

Yet, two intriguing problems remain with this explanation. First, if we extend the same rationale to comparisons between sexuality and asexuality, we would be expected to be unable to detect any difference in rates of protein evolution when comparing these two breeding systems. Asexual lineages, which are often of recent origin (Law and Crespi 2002; Neiman et al. 2005; Johnson 2006; Paland and Lynch 2006), would be more prone to accumulate deleterious mutations and could become extinct as fast, if not faster, as selfing species. However, recent studies have shown that recent transitions from sexual to asexual lineages result in detectable excess of amino acid substitutions, at least in mitochondrial genes, in *Daphnia* (Paland and Lynch 2006) and *Campeloma* freshwater snails (Johnson and Howard 2007). Second, if selfing lineages become extinct before the accumulation of deleterious mutations being detectable, what is the very cause of their extinction? Beyond the observed levels of diversity, the hypothesis of limited adaptive potential in selfers remains largely unexplored. Additional studies in other groups are crucially needed to compare the potential for adaptation between selfers and outcrossers and to confirm or reject the lack of signatures of accumulation of slightly deleterious mutations.

ACKNOWLEDGMENTS

We thank N. Galtier, G. Marais, and L. Duret for stimulating discussions, C. Scornavacca for congruence analyses, two anonymous reviewers for comments, and B. Ramora and S. Santoni for technical help. JSE and AC were founded by the Institut National de la Recherche Agronomique (INRA), postdoctoral grant program. This work was supported by the INRA, Tritipol initiative, and Agence Nationale de la Recherche, Exegese program (grant number ANR-05-BLANC-0258-01) and CoGeBi program (grant number ANR-08-GENM-036-01). *Brachypodium distachyon* sequences were produced by the US Department of Energy Joint Genome Institute (http://www.jgi.doe.gov/). This publication is the contribution ISEM 2010-027 of the Institut des Sciences de l'Evolution de Montpellier (UMR 5554 – CNRS).

LITERATURE CITED

Abdel-Ghani, A. H., H. K. Parzies, A. Omary, and H. H. Geiger. 2004. Estimating the outcrossing rate of barley landraces and wild barley populations collected from ecologically different regions of Jordan. Theor. Appl. Genet. 109:588–595.

Akhunov, E. D., A. W. Goodyear, S. Geng, L. L. Qi, B. Echalier, B. S. Gill, Miftahudin, J. P. Gustafson, G. Lazo, S. M. Chao, et al. 2003. The organization and rate of evolution of wheat genomes are correlated with recombination rates along chromosome arms. Genome Res. 13:753–763.

Armbruster, W. S. 1993. Evolution of plant pollination systems: hypotheses and tests with the neotropical vine *Dalechampia*. Evolution 47:1480–1505.

Baker, H. G. 1955. Self-incompatibility and establishment after "long-distance" dispersal. Evolution 9:347–348.

———. 1967. Support for Baker's law as a rule. Evolution 21:853–856.

Barrett, S. C. H., L. D. Harder, and A. C. Worley. 1996a. Comparative biology of plant reproductive traits. Philos. Trans. R. Soc. Lond. B 351:1272–1280.

———. 1996b. The comparative biology of pollination and mating in flowering plants. Philos. Trans. R. Soc. Lond. B 351:1271–1280.

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

Baumann, U., J. Juttner, X. Bian, and P. Langridge. 2000. Self-incompatibility in the grasses. Ann. Bot. 85:203–209.

Bena, G., B. Lejeune, J.-M. Prosperi, and I. Olivieri. 1998. Molecular phylogenetic approach for studying life-history evolution: the ambiguous example of the genus *Medicago* L. Proc. R. Soc. Lond. B 265:1141–1151.

Betancourt, A. J., J. J. Welch, and B. Charlesworth. 2009. Reduced effectiveness of selection caused by a lack of recombination. Curr. Biol. 19:655–660.

Birdsell, J. A. 2002. Integrating genomics, bioinformatics, and classical genetics to study the effects of recombination on genome evolution. Mol. Biol. Evol. 19:1181–1197.

Blattner, F. R. 2004. Phylogenetic analysis of *Hordeum* (Poaceae) as inferred by nuclear rDNA ITS sequences. Mol. Phylogenet. Evol. 33:289–299.

Cenci, A., S. Somma, N. Chantret, J. Dubcovsky, and A. Blanco. 2004. PCR identification of durum wheat BAC clones containing genes coding for carotenoid biosynthesis enzymes and their chromosome localization. Genome 47:911–917.

Chantret, N., J. Salse, F. Sabot, S. Rahman, A. Bellec, B. Laubin, I. Dubois, C. Dossat, P. Sourdille, P. Joudrier, et al. 2005. Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (*Triticum* and *Aegilops*). Plant Cell 17:1033–1045.

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

———. 2009. Effective population size and patterns of molecular evolution and variation. Nat. Rev. Genet. 10:195–205.

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

Charlesworth, D. 2006. Evolution of plant breeding systems. Curr. Biol. 16:R726–R735.

Charlesworth, D., and B. Charlesworth. 1987. Inbreeding depression and its evolutionary consequences. Annu. Rev. Ecol. Evol. Syst. 18:237–268.

———. 1995. Quantitative genetics in plants: the effect of the breeding system on genetic variability. Evolution 49:911–920.

Charlesworth, D., and S. I. Wright. 2001. Breeding systems and genome evolution. Curr. Opin. Genet. Develop. 11:685–690.

Charlesworth, D., B. Charlesworth, and M. T. Morgan. 1995. The pattern of neutral molecular variation under the background selection model. Genetics 141:1619–1632.

Cruden, R. W. 1977. Pollen-ovule ratios: a conservative indicator of breeding systems in flowering plants. Evolution 31:32–46.

———. 2000. Pollen grains: why so many? Plant Syst. Evol. 222:143–165.

Cutter, A. D., J. D. Wasmuth, and N. L. Washington. 2008. Patterns of molecular evolution in *Caenorhabditis* preclude ancient origins of selfing. Genetics 178:2093–2104.

Damgaard, C., and V. Loeschcke. 1994. Genotypic variation for reproductive characters, and the influence of pollen-ovule ratio on selfing rate in rape seed (*Brassica napus*). J. Evol. Biol. 7:599–607.

Darwin, C. 1876. The different forms of flowers on plants of the same species. John Murray, London.

———. 1878. The effects of cross and self-fertilization in the vegetal kingdom. John Murray, London.

Depace, C., and C. O. Qualset. 1995. Mating system and genetic differentiation in *Dasypyrum villosum* (Poaceae) in Italy. Plant Syst. Evol. 197:123–147.

Dubcovsky, J., M. C. Luo, G. Y. Zhong, R. Bransteitter, A. Desai, A. Kilian, A. Kleinhofs, and J. Dvorak. 1996. Genetic map of diploid wheat, *Triticum monococcum* L, and its comparison with maps of *Hordeum vulgare* L. Genetics 143:983–999.

Duret, L., and P. F. Arndt. 2008. The impact of recombination on nucleotide substitutions in the human genome. PLoS Genet. 4:e1000071.

Duret, L., and N. Galtier. 2009. Biased gene conversion and the evolution of mammalian genomic landscapes. Annu. Rev. Genom. Human Genet. 10:285–311.

Dutheil, J., and B. Boussau. 2008. Non-homogeneous models of sequence evolution in the Bio++ suite of libraries and programs. BMC Evol. Biol. 8:255.

Dvorak, J., and E. D. Akhunov. 2005. Tempos of gene locus deletions and duplications and their relationship to recombination rate during diploid and polyploid evolution in the Aegilops-Triticum alliance. Genetics 171:323–332.

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

Ferrer, M. M., and S. V. Good-Avila. 2006. Macrophylogenetic analyses of the gain and loss of self-incompatibility in the Asteraceae. New Phytol. 173:401–414.

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

Foxe, J. P., T. Slotte, E. A. Stahl, B. Neuffer, H. Hurka, and S. W. Wright. 2009. Recent speciation associated with the evolution of selfing in *Capsella*. Proc. Natl. Acad. Sci. USA 106:5241–5245.

Galtier, N., and L. Duret. 2007. Adaptation or biased gene conversion? Extending the null hypothesis of molecular evolution. Trends Genet. 23:273–277.

Galtier, N., G. Piganeau, D. Mouchiroud, and L. Duret. 2001. GC-content evolution in mammalian genomes: the biased gene conversion hypothesis. Genetics 159:907–911.

Galtier, N., L. Duret, S. Glemin, and V. Ranwez. 2009. GC-biased gene conversion promotes the fixation of deleterious amino acid changes in primates. Trends Genet. 25:1–5.

Glémin, S. 2007. Mating systems and the efficacy of selection at the molecular level. Genetics 177:905–916.

Glémin, S., E. Bazin, and D. Charlesworth. 2006. Impact of mating systems on patterns of sequence polymorphism in flowering plants. Proc. R. Soc. Lond. B 273:3011–3019.

Goldberg, E. E., and B. Igic. 2008. On phylogenetic tests of irreversible evolution. Evolution 62:2727–2741.

Goodwillie, C. 1997. The genetic control of self-incompatibility in *Linanthus parviflorus* (Polemoniaceae). Heredity 79:424–432.

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

Goodwillie, C., S. Kalisz, and C. G. Eckert. 2005. The evolutionary enigma of mixed mating in plants: occurrence, theoretical explanations, and empirical evidence. Ann. Rev. Ecol. Evol. Syst. 36:47–79.

Guo, Y.-L., J. S. Bechsgaard, T. Slotte, B. Neuffer, M. Lascoux, D. Weigel, and M. H. Schierup. 2009. Recent speciation of *Capsella rubella* from

Capsella grandiflora, associated with loss of self-incompatibility and an extreme bottleneck. Proc. Natl. Acad. Sci. USA 106:5246–5251.

Haddrill, P. R., D. L. Halligan, D. Tomaras, and B. Charlesworth. 2007. Reduced efficacy of selection in regions of the *Drosophila* genome that lack crossing over. Genome Biol. 8:R18.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Series 41:95–98.

Hamrick, J. L., and M. J. W. Godt. 1996. Effects of life history traits on genetic diversity in plants species. Philos. Trans. R. Soc. Lond. B 351:1291–1298.

Haudry, A., A. Cenci, C. Guilhaumon, E. Paux, S. Poirier, S. Santoni, J. David, and S. Glémin. 2008. Mating system and recombination affect molecular evolution in four Triticeae species. Genet. Res. 90:97–109.

Hegde, S. G., J. Valkoun, and J. G. Waines. 2000. Genetic diversity in wild wheats and goat grass. Theor. Appl. Genet. 101:309–316.

Huelsenbeck, J. P., and F. Ronquist. 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Huo, N. X., Y. Q. Gu, G. R. Lazo, J. P. Vogel, D. Coleman-Derr, M. C. Luo, R. Thilmony, D. F. Garvin, and O. D. Anderson. 2006. Construction and characterization of two BAC libraries from *Brachypodium distachyon*, a new model for grass genomics. Genome 49:1099–1108.

Igic, B., L. Bohs, and J. R. Kohn. 2006. Ancient polymorphism reveals unidirectional breeding system shifts. Proc. Natl. Acad. Sci. USA 103:1359–1363.

Ingvarsson, P. K. 2002. A metapopulation perspective on genetic diversity and differentiation in partially self-fertilizing plants. Evolution 56:2368–2373.

Johnson, S. G. 2006. Geographic ranges, population structure and ages of sexual and parthenogenetic snail lineages. Evolution 60:1417–1426.

Johnson, S. G., and R. S. Howard. 2007. Contrasting patterns of synonymous and nonsynonymous sequence evolution in asexual and sexual freshwater snail lineages. Evolution 61:2728–2735.

Kahler, A. L., M. T. Clegg, and R. W. Allard. 1975. Evolutionary changes in the mating system of an experimental population of barley (*Hordeum vulgare* L.). Proc. Natl. Acad. Sci. USA 72:943–946.

Kaplan, N. L., R. R. Hudson, and C. H. Langley. 1989. The "hitchhiking effect" revisited. Genetics 123:887–899.

Kellogg, E. A., R. Appels, and R. J. Mason-Gamer. 1996. When genes tell different stories: the diploid genera of Triticeae (Gramineae). Syst. Bot. 21:321–347.

Kiontke, K., N. P. Gavin, Y. Raynes, C. Roehrig, F. Piano, and D. H. A. Fitch. 2004. *Caenorhabditis* phylogeny predicts convergence of hermaphroditism and extensive intron loss. Proc. Natl. Acad. Sci. USA 101:9003–9008.

Kohn, J. R., S. W. Graham, B. Morton, J. J. Doyle, and S. C. H. Barrett. 1996. Reconstruction of the evolution of reproductive characters in Pontederiaceae using phylogenetic evidence from chloroplast DNA restriction-site variation. Evolution 50:1454–1469.

Lande, R., and D. Schemske. 1985. The evolution of self-fertilization and inbreeding depression in plants. I. Genetics models. Evolution 39:24–40.

Lartillot, N., and H. Philippe. 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Mol. Biol. Evol. 21:1095–1109.

Law, J. H., and B. J. Crespi. 2002. Recent and ancient asexuality in *Timema* walkingsticks. Evolution 56:1711–1717.

Luo, M. C., Z. L. Yang, R. S. Kota, and J. Dvorak. 2000. Recombination of chromosomes 3A(m) and 5A(m) of *Triticum monococcum* with homeologous chromosomes 3A and 5A of wheat: the distribution of recombination across chromosomes. Genetics 154:1301–1308.

Luo, M. C., K. R. Deal, Z. L. Yang, and J. Dvorak. 2005. Comparative genetic maps reveal extreme crossover localization in the *Aegilops speltoides* chromosomes. Theor. Appl. Genet. 111:1098–1106.

Lynch, M., J. Conery, and R. Bürger. 1995. Mutation accumulation and the extinction of small populations. Am. Nat. 146:489–518.

Maddison, W. P., and D. R. Maddison. 2008. Mesquite: A modular system for evolutionary analysis. Version 2.5.

Marais, G. 2003. Biased gene conversion: implications for genome and sex evolution. Trends Genet. 19:330–338.

Marais, G., B. Charlesworth, and S. I. Wright. 2004. Recombination and base composition: the case of the highly self-fertilizing plant *Arabidopsis thaliana*. Genome Biol. 5:R45.

Mason-Gamer, R. J. 2005. The β-amylase genes of grasses and a phylogenetic analysis of the Triticeae (Poaceae). Am. J. Bot. 92:1045–1058.

Michalski, S. G., and W. Durka. 2009. Pollination mode and life form strongly affect the relation between mating system and pollen to ovule ratios. New Phytol. 183:470–479.

Morran, L. T., B. J. Cappy, J. L. Anderson, and P. C. Phillips. 2009. Sexual partners for the stressed: facultative outcrossing in the self-fertilizing nematode *Caenorhabditis elegans*. Evolution 63:1473–1482.

Morrell, P. L., D. M. Toleno, K. E. Lundy, and M. T. Clegg. 2005. Low levels of linkage disequilibrium in wild barley (*Hordeum vulgare* ssp. *spontaneum*) despite high rates of self-fertilization. Proc. Natl. Acad. Sci. USA 102:2442–2447.

Munkvold, J. D., R. A. Greene, C. E. Bertmudez-Kandianis, C. M. La Rota, H. Edwards, S. F. Sorrells, T. Dake, D. Benscher, R. Kantety, A. M. Linkiewicz, et al. 2004. Group 3 chromosome bin maps of wheat and their relationship to rice chromosome 1. Genetics 168:639–650.

Neiman, M., J. Jokela, and C. M. Lively. 2005. Variation in asexual lineage age in *Potamopyrgus antipodarum*, a New Zealand snail. Evolution 59:1945–1952.

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

Nordborg, M., B. Charlesworth, and D. Charlesworth. 1996. The effect of recombination on background selection. Genet. Res. 67:159–174.

Nybom, H. 2004. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Mol. Ecol. 13:1143–1155.

Pagel, M. 1999. Inferring the historical patterns of biological evolution. Nature 401:877–884.

Pagel, M., A. Meade, and D. Barker. 2004. Bayesian estimation of ancestral character states on phylogenies. Syst. Biol. 53:673–684.

Paland, S., and M. Lynch. 2006. Transitions to asexuality result in excess amino acid substitutions. Science 311:990–992.

Parzies, H. K., W. Spoor, and R. A. Ennos. 2000. Genetic diversity of barley landrace accessions (*Hordeum vulgare* ssp. *vulgare*) conserved for different lengths of time in *ex situ* gene banks. Heredity 84:476–486.

Petersen, G., and O. Seberg. 1997. Phylogenetic analysis of the Triticeae (Poaceae) based on *rpo*A sequence data. Mol. Phylogenet. Evol. 7:217–230.

Polanco, C., C. Gonzalez, F. J. Vences, and M. P. Delavega. 1994. Non-random mating in a *Secale cereale* L (Rye) population. Heredity 72:549–556.

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

Posada, D., and K. A. Crandall. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14:817–818.

Presgraves, D. C. 2005. Recombination enhances protein adaptation in *Drosophila melanogaster*. Curr. Biol. 15:1651–1656.

Rasband, W. 2007. ImageJ 1.34s. Pp. image processing and analysis in Java. National Institutes of Health, USA.

Saintenac, C., M. Falque, O. C. Martin, E. Paux, C. Feuillet, and P. Sourdille. 2009. Detailed recombination studies along chromosome 3B provide new insights on crossover distribution in wheat (*Triticum aestivum* L.). Genetics 181:393–403.

Sanderson, M. J. 1993. Reversibility in evolution: a maximum likelihood approach to character gain/loss bias in phylogenies. Evolution 47:236–252.

Savolainen, O., C. H. Langley, B. P. Lazzaro, and H. Fréville. 2000. Contrasting patterns of nucleotide polymorphism at the alcohol dehydrogenase locus in the outcrossing *Arabidopsis lyrata* and the selfing *Arabidopsis thaliana*. Mol. Biol. Evol. 17:645–655.

Schemske, D., and R. Lande. 1985. The evolution of self-fertilization and inbreeding depression in plants. II. Empirical observations. Evolution 39:41–52.

Schoen, D. 1982. Male reproductive effort and breeding system in an hermaphroditic plant. Oecologia 53:255–257.

Schoen, D., M. Johnston, and A. L'Heureux. 1997. Evolutionary history of the mating system in *Amsinckia* (Boraginaceae). Evolution 51:1090–1099.

Seberg, O., and S. Frederiksen. 2001. A phylogenetic analysis of the monogenomic Triticeae (Poaceae) based on morphology. Bot. J. Linn. Soc. 136:75–97.

Sorrells, M. E., M. La Rota, C. E. Bermudez-Kandianis, R. A. Greene, R. Kantety, J. D. Munkvold, Miftahudin, A. Mahmoud, X. Ma, P. J. Gustafson, et al. 2003. Comparative DNA sequence analysis of wheat and rice genomes. Genome Res. 13:1818–1827.

Staden, R., D. P. Judge, and J. K. Bonfield. 2000. The Staden package, 1998. Methods Mol. Biol. 132:115–130.

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

Sueoka, N. 1962. On the genetic basis of variation and heterogeneity of DNA base composition. Proc. Natl. Acad. Sci. USA 48:582–592.

Swofford, D. L. 2003. PAUP*: Phylogenetic analysis using parsimony (* and other methods), version 4.0b10. Sinauer Associates, Sunderland, MA.

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

Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24:1596–1599.

Tang, C., C. Toomajian, S. Sherman-Broyles, V. Plagnol, Y.-L. Guo, T. T. Hu, R. M. Clark, J. B. Nasrallah, D. Weigel, and M. Nordborg. 2007. The evolution of selfing in *Arabidopsis thaliana*. Science 317:1070–1072.

Verhoeven, K. J. F., K. L. Simonsen, and L. M. McIntyre. 2005. Implementing false discovery rate control: increasing your power. Oikos 108:643–647.

von Bothmer, R., N. Jacobsen, C. J. Baden, R. B. Jorgensen, and I. Linde-Laursen. 1995. An ecogeographical study of the genus *Hordeum*. International Plant Genetic Ressources Institute, Rome.

Webster, M. T., E. Axelsson, and H. Ellegren. 2006. Strong regional biases in nucleotide substitution in the chicken genome. Mol. Biol. Evol. 23:1203–1216.

Wright, S. I., B. Lauga, and D. Charlesworth. 2002. Rates and patterns of molecular evolution in inbred and outbred *Arabidopsis*. Mol. Biol. Evol. 19:1407–1420.

Wright, S. I., J. P. Foxe, L. DeRose-Wilson, A. Kawabe, M. Looseley, B. S. Gaut, and D. Charlesworth. 2006. Testing for effects of recombination rate on nucleotide diversity in natural populations of *Arabidopsis lyrata*. Genetics 174:1421–1430.

Wright, S. I., G. Iorgovan, S. Misra, and M. Mokhtari. 2007. Neutral evolution of synonymous base composition in the Brassicaceae. J. Mol. Evol. 64:136–141.

Yang, Y., J. Han, and J. D. Berdahl. 2008. Allozyme diversity of Russian wildrye accessions. Plant Breeding 127:480–484.

Yang, Z. 2007. PAML 4: a program package for phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24:1586–1591.

Yang, Z., W. S. Wong, and R. Nielsen. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. Mol. Biol. Evol. 22:1107–1118.

Zaharieva, M., and P. Monneveux. 2006. Spontaneous hybridization between bread wheat (*Triticum aestivum* L.) and its wild relatives in Europe. Crop Science 46:512–527.

Zhang, H., S. M. Reader, X. Liu, J. Z. Jia, M. D. Gale, and K. M. Devos. 2001. Comparative genetic analysis of the *Aegilops longissima* and *Ae. sharonensis* genomes with common wheat. Theor. Appl. Genetics 103:518–525.

Zohary, D., and D. Imber. 1963. Genetic dimorphism in fruit types in *Aegilops speltoides*. Heredity 18:223–231.

Associate Editor: J. Pannell

Supporting Information

The following supporting information is available for this article:

Figure S1. Hierarchical models of sequences evolution.

Table S1. Species, accession numbers in the US Department of Agriculture database, and geographic origin of Triticeae.

Table S2. Individual gene topologies.

Table S3. Traits related to the mating system measured in the studied species.

Table S4. Effect of the mating system on substitution parameters when controlling for the life cycle (annual vs. perennial), when using the multigenic phylogeny instead of the gene phylogeny, and when internal branches are declared as outcrossing (i.e., all transitions are from outcrossing to selfing).

Supporting Information may be found in the online version of this article.