Letters

A re-evaluation of the homoploid hybrid origin of *Aegilops tauschii*, the donor of the wheat D-subgenome

Introduction

Hybridization is a prominent evolutionary force promoting plant diversification, either with or without subsequent genome doubling (Abbott *et al.*, 2013; Soltis *et al.*, 2014; Yakimowski & Rieseberg, 2014). The *Aegilops–Triticum* complex is an ideal system to investigate how natural hybridization and allopolyploidization have caused species diversification (Matsuoka, 2011). Recently, Marcussen *et al.* (2014) proposed the tantalizing scenario that the ancestral D lineage originated via homoploid hybridization between ancient A and B lineages some 5 million years ago (Mya) (the definition of A, B and D lineages shown in Fig. 1). Evidence for this mode of origin was derived from phylogenomic and population genetic analyses of nuclear genes, but without taking into account the evolutionary history and chloroplast topology of this species complex. Meanwhile, in a recent issue of *New Phytologist*, Gornicki *et al.* (2014) reported the chloroplast phylogeny of the *Triticum–Aegilops* complex based on 25 chloroplast genomes of eight modern A, S and D genome diploid species and four polyploid wheat species, but they did not address the origin of the D genome. Here, by re-analyzing critical data used by both studies and additional data, we present evidence for a more complex hybrid origin of the D genome of *A. tauschii*.

Evolutionary history of *Triticum–Aegilops* complex

To date, 13 diploid species of the *Triticum–Aegilops* complex, which belong to eight distinct but related genome groups (A, D, S, M, C, U, N and T), have been identified (Table 1) and which are variously and sometimes sympatrically distributed in the Middle East (Lilienfeld, 1951; Gill & Friebe, 2002; Huang *et al.*, 2002). The A, B and D genomes, harbored by diploid species *T. urartu*, *A. speltoides* (or a closely related species) and *A. tauschii*, respectively, are established as the diploid genome donors of the A-, B- and D-subgenomes of hexaploid bread wheat, *T. aestivum* (Cox, 1998; Huang *et al.*, 2002; Petersen *et al.*, 2006). The remaining genomes harbored by the diploid species (except the T genome) are found in polyploid *Aegilops* species (Gill & Friebe, 2002). Eighteen naturally occurring allopolyploid species have been described (Table 1), which also are widely distributed across the Near East. On the basis of plant habit, spike morphology and cytological data,

Zohary & Feldman (1962) classified these allopolyploid species into three major genome groups (called cytological clusters), which are A, U and D (Table 1). Variation patterns of the three cytological clusters led to the hypothesis that the current allopolyploid species within each cluster probably evolved from only a few initial amphidiploids (Zohary & Feldman, 1962; Feldman, 1965; Pazy & Zohary, 1965). This hypothesis is further supported by cytological and phylogenetic analyses showing that the A-, U- and D-genomes have indeed donated the maternal genome to most of the current polyploid species (Kimber & Tsunewaki, 1988; Meimberg *et al.*, 2009; Tsunewaki, 2009). Together, these attributes suggest that hybridization, either at the homoploid level or followed by polyploidization, has occurred frequently within the *Triticum–Aegilops* complex.

Homoploid hybridization origin of *Aegilops tauschii*

Marcussen *et al.* (2014) proposed that all extant diploid species of the *Aegilops–Triticum* complex are derived from A, B and D lineages (Fig. S6 in Marcussen *et al.*, 2014). Phylogenomic analyses based on nuclear genome sequences revealed that the phylogenetic positions of A (*T. monococcum/ T. urartu*), B (*A. speltoides*) and D (*A. tauschii*) genome species varied among nuclear genes, with topologies A (B, D) and B (A, D) each being about twice as common as D (A, B) (Table 1 in Marcussen *et al.*, 2014). However, evidence for the homoploid hybrid origin of *A. tauschii* derives from phylogenomic analyses of modern S, A and D genomes, without taking into account the other genomes (e.g. M, N, T, U and C) within this species complex. Given that the breadth of taxonomic sampling could affect the identification of hybridization events, we re-analyzed the topologies of the four focal genomes (A, S, D and S^{sh}) for the 275 nuclear genes used in Marcussen *et al.* (2014), who proposed a homoploid hybrid origin of the D lineage (Table S4 in Marcussen *et al.*, 2014). If the D lineage species were formed from a single homoploid hybridization event, as proposed by Marcussen *et al.* (2014), then all species derived from this event would be expected to cluster together at most nuclear genes. However, our results showed that the two D lineage species, *A. tauschii* and *A. sharonensis*, are separated in 40% of the 209 gene trees (Supporting Information Fig. S1). Marcussen *et al.* (2014) also reported a distinct hybrid pattern for *A. sharonensis*, with only *c.* 25% of the 275 gene trees reflecting B-lineage ancestry. Indeed, Waines & Johnson (1972) have documented that *A. sharonensis* is a hybrid between *A. lon*gissima and *A. bicornis*. In addition, cytogenetic analyses revealed that modern S* genome species (D lineage) are closer to *A. speltoides* (B lineage) than to *A. tauschii* (D lineage) (Kihara, 1954). Based on these observations, we hypothesize that the origin of *A. tauschii* may be more complicated than envisioned by Marcussen *et al.* (2014).
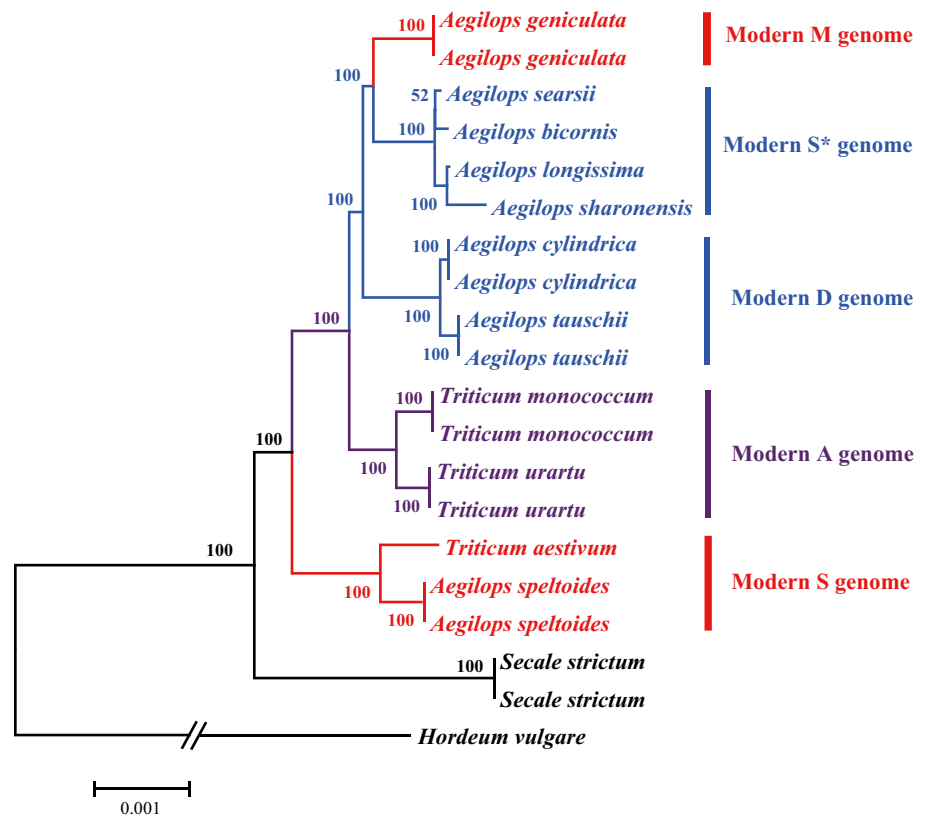
Fig. 1 Neighbor joining tree of the *Triticum–Aegilops* complex based on whole chloroplast genome sequences. S* includes the four *Sitopsis* species *A. bicornis* (S^b), *A. searsii* (S^s), *A. sharonensis* (S^sh) and *A. longissima* (S^l), whereas S indicates *A. speltoides*. Red, purple and blue colors represent S-, A- and D-lineages, respectively.

Table 1 Species of *Triticum–Aegilops* complex and their genomic constitution, lineage and group

Diploid species	Chloroplast	Nuclear	Allopolyploid species	Chloroplast	Nuclear	Group
Aegilops tauschii	D	DD	*Aegilops neglecta*	U	UUMM	U
Aegilops sharonensis	S^sh	S^shS^sh	*Aegilops neglecta* var. *recta*	U	UUMMNN	U
Aegilops bicornis	S^b	S^bS^b	*Aegilops biuncialis*	U	UUMM	U
Aegilops longissima	S^l	S^lS^l	*Aegilops columnaris*	U	UUMM	U
Aegilops searsii	S^s	S^sS^s	*Aegilops geniculata*	M	UUMM	U
Aegilops speltoides	S	SS	*Aegilops kotschyi*	S^v	UUS^vS^v	U
Aegilops caudata	C	CC	*Aegilops peregrina*	S^v	UUS^vS^v	U
Aegilops uniaristata	N	NN	*Aegilops triuncialis*	U/C	UUCC	U
Aegilops comosa	M	MM	*Aegilops ventricosa*	D	DDNN	D
Aegilops umbellulata	U	UU	*Aegilops cylindrica*	D	DDCC	D
Triticum urartu	A	AA	*Aegilops vavilovii*	D	DDMMSS	D
Triticum monococcum	A	AA	*Aegilops juvenalis*	D	DDMMUU	D
Amblyopyrum mutica	T	TT	*Aegilops crassa*	D	DDMM	D
			Aegilops crassa var. *glumiaristata*	na	DDDDMM	D
			Triticum turgidum	S	AABB	A
			Triticum timopheevii	S	AAGG	A
			Triticum aestivum	S	AABBDD	A
			Triticum zhukovskyi	A	AAAAGG	A

Information of genomic constitution is retrieved from Gill & Friebe (2002) with minor revision. The subdivision of D, M and U genomes are not shown. The definition of genome group for each allopolyploid species is defined according to Zohary & Feldman (1962). na, no related information.

Chloroplast view of the origin of *Aegilops tauschii*

Additional information bearing on the history of the *Triticum–Aegilops* complex may derive from analysis of chloroplast genomes, which are maternally inherited in this species complex (Fukasawa, 1959; Kihara, 1959). The cpDNA topology of this group has been investigated in previous studies using selected genes (Hirai & Tsunewaki, 1981; Terachi *et al.*, 1987; Wang *et al.*, 1997; Yamane & Kawahara, 2005; Meimberg *et al.*, 2009), and more recently whole genome sequences (Gornicki *et al.*, 2014; Middleton *et al.*, 2014). These studies focused on the origin of domesticated wheat and phylogenetic relationships within the *Triticum–Aegilops* complex. For example, Gornicki *et al.* (2014) revealed a basal clade of *A. speltoides* and *A. tauschii* grouped

together with four S* genome species (*A. longissima*, *A. searsii*, *A. sharonensis* and *A. bicornis*). However, the other diploid genome (e.g. M genome) and polyploid species within *Triticum–Aegilops* complex were not included in their study. Middleton *et al.* (2014) also reported a similar chloroplast topology with five A, S, D diploid species and three polyploid species (including M genome), but the S* genome species were not included in their study. To encompass all diploid and polyploid genomes in a single framework, we retrieved 20 whole chloroplast genome sequences from GenBank deposited by Gornicki *et al.* (2014) and Middleton *et al.* (2014) and conducted neighbor-joining phylogenetic analysis using MEGA 6 (Tamura *et al.*, 2013) (Fig. 1). The resulting topology showed that the chloroplast genome of *A. tauschii* (D genome) is more closely related to that of other D and all S* genome species (including *A. sharonensis*) than it is to either the *A. speltoides* (S genome) or *T. monococcum/T. urartu* (A genome) chloroplast genomes. This result would not be expected under the scenario of a single homoploid hybrid origin of *A. tauschii*, in which case the cpDNA of *A. tauschii* would be expected to be more similar to that of either the A or B genome lineages. Marcussen *et al.* (2014) proposed that *A. tauschii* originated from homoploid hybridization between *A. speltoides* and the ancestor of modern *T. urartu/ T. monococcum*. If this were the case, *A. tauschii* would be expected to share a chloroplast genome with one of these two putative progenitor lineages. However, the chloroplast topology reveals that *A. tauschii* is cladistically nested between the A and remaining S* and M genomes (Fig. 1).

One possible explanation for the apparent discrepancy between the cpDNA-derived phylogeny and the interpretation of Marcussen *et al.* (2014) is that *A. speltoides* itself underwent later hybridization(s), in which it captured the chloroplast

Table 2 Number of single nucleotide polymorphisms (SNPs) shared among A, S (~B), M, S* and D genomes based on the whole chloroplast genome

	S genome	A genome	D genome	S* genome
A genome	224			
D genome	191	306		
S* genome	192	287	295	
M genome	182	272	283	306
Ancient D lineage	56	149	na	na

S* genome includes the four *Sitopsis* species *Aegilops bicornis* (S^b genome), *A. searsii* (S^s genome), *A. sharonensis* (S^{sh} genome) and *A. longissima* (S^l genome); D genome lineage contains *A. geniculata* (M genome), *A. tauschii* (D genome) and the four *Sitopsis* species; na, data is unavailable.

genome from some other species in the B genome lineage. As shown in the chloroplast topology, however, *A. speltoides* is basal in the clade, differing phylogenetically and quantitatively from the remaining species in the *Triticum–Aegilops* complex. The latter is evidenced in an analysis of single nucleotide polymorphisms (SNPs), in which *A. speltoides* shares fewer SNPs with the remaining genomes than they do with each other (Table 2). These attributes indicate that *A. speltoides* possesses a rather distinct chloroplast genome. In addition, the number of shared SNPs between D and A genomes is only slightly higher than between D and M or between D and S* genomes (Table 2), suggesting that the D genome is equidistant from the A, S* and M genomes. An alternative scenario is that the origin of the D genome lineage that gave rise to *A. tauschii* is more complex than that proposed by Marcussen *et al.* (2014). For example, *A. speltoides* and *T. urartu/ T. monococcum* could have evolved from the ancient B and A genomes, respectively, with the remaining species (including
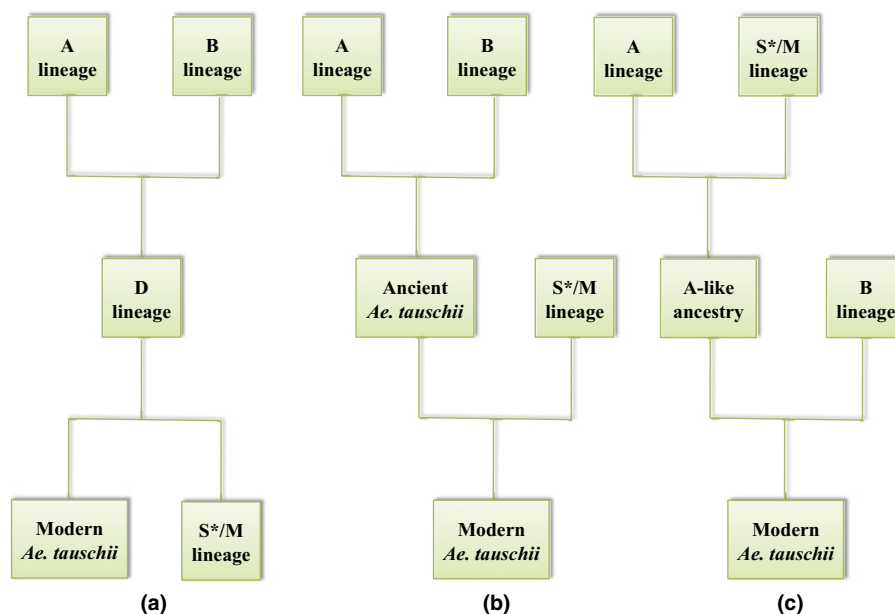

Fig. 2 Three models of homoploid hybridization origin for *Aegilops tauschii*. (a) Direct homoploid hybridization model; (b) chloroplast capture model; (c) ancestry capture model. The A, B and S*/M lineages include the modern A, S and S*/M genomes, respectively; D lineage includes the modern D, S* and M genomes. See the 'Chloroplast view of the origin of *Aegilops tauschii*' section for details.

the modern D, S* and M genomes in Fig. 1) of *Triticum–Aegilops* complex derived from the ancient D lineage (Fig. 2a). If the ancient D lineage were to have evolved through a single homoploid hybridization between A and B lineages, it would be expected to phylogenetically cluster with either of its two donors and have a cpDNA genome that is much more like one parent than the other. The chloroplast topology reveals that the ancient D lineage clusters together with A lineage (Fig. 1). Indeed, the number of shared SNPs between the A and D lineages is higher than that between the B and D lineages (Table 2). Under this hypothesis, recent hybridization might be responsible for the distinct hybrid pattern of modern *A. tauschii* and *A. sharonensis* in the phylogenomic analyses.

Another possibility for the origin of modern *A. tauschii* is a chloroplast capture model (Fig. 2b). Under this scenario A and B lineages were involved in homoploid hybridization, leading to the formation of an *A. tauschii*-like nuclear genome (ancestry of modern *A. tauschii*), which was followed by hybridization(s) with a third species in the S*/M genomes. In this case, modern *A. tauschii* captured the chloroplast genome from S*/M genome species without much nuclear introgression. Alternatively, modern *A. tauschii* might have originated from homoploid hybridization between the A and B lineages, but the hybridization event was preceded by an earlier hybridization with a third species that donated the S*/M lineage chloroplast genome to the maternal parent of the later homoploid hybridization event (Fig. 2c). If this were the case, modern *A. tauschii* would have evolved with A and B lineage nuclear genomes, but with a S*/M type chloroplast genome.

Taken together, our integrated re-evaluation, while confirming the hybrid nature of *A. tauschii*, points clearly to a more complex history of the species than that proposed by Marcussen *et al.* (2014), one that may have involved multiple rounds of both recent and ancient hybridizations.

Chloroplast genome data of *Triticum–Aegilops* complex

All chloroplast genome sequences deposited by Gornicki *et al.* (2014) and Middleton *et al.* (2014) were retrieved from GenBank. *Aegilops speltoides*, JQ740834, NC_022135; *A. sharonensis*, KJ614419; *A. bicornis*, KJ614417; *A. longissima*, KJ614416; *A. searsii*, KJ614415; *A. tauschii*, JQ754651, NC_022133; *Triticum monococcum*, KC912690, NC_021760; *T. urartu*, KC912693, NC_021762; *T. aestivum*, KC912694; *A. cylindrica*, KF534489, NC_023096; *A. geniculata*, KF534490, NC_023097; *Hordeum vulgare*, KC912687; *Secale cereale*, KC912691, NC_021761.

Acknowledgements

Lin-Feng Li[1,2]*, Bao Liu[1], Kenneth M. Olsen[2] and Jonathan F. Wendel[3]

[1]Key Laboratory of Molecular Epigenetics of Ministry of Education, Northeast Normal University, No. 5268 Renmin Str., Changchun 130024, China;
[2]Department of Biology, Washington University in St Louis, St Louis, MO 63130, USA;
[3]Department of Ecology, Evolution and Organismal Biology, Iowa State University, Ames, IA 50011, USA
(*Author for correspondence: tel +86 431 8509 9367; email lilf241@nenu.edu.cn)

References

Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJE, Bierne N, Boughman J, Brelsford A, Buerkle CA, Buggs R *et al.* 2013. Hybridization and speciation. *Journal of Evolutionary Biology* 26: 229–246.

Cox TS. 1998. Deepening the wheat gene pool. *Journal of Crop Production* 1: 1–25.

Feldman M. 1965. Further evidence for natural hybridization between tetraploid species of *Aegilops* section *Pleionathera. Evolution* 19: 562–574.

Fukasawa H. 1959. Nucleus substitution and restoration by means of successive backcrosses in wheat and its related genus *Aegilops. Journal of Japanese Botany* 17: 55–91.

Gill BS, Friebe B. 2002. Cytogenetics, phylogeny and evolution of cultivated wheats. In: Curtis BC, Rajaram S, Gómez Macpherson H, eds. *Bread wheat: improvement and production.* Rome, Italy: Food and Agriculture Organization of the United Nations, 71–88.

Gornicki P, Zhu H, Wang J, Challa GS, Zhang Z, Gill BS, Li W. 2014. The chloroplast view of the evolution of polyploid wheat. *New Phytologist* 204: 704–714.

Hirai A, Tsunewaki K. 1981. Genetic diversity of the cytoplasm in *Triticum* and *Aegilops.* VIII. Fraction I protein of 39 cytoplasms. *Genetics* 99: 487–493.

Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R, Gornicki P. 2002. Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. *Proceedings of the National Academy of Sciences, USA* 99: 8133–8138.

Kihara H. 1954. Consideration on the evolution and distribution of *Aegilops* species based on the analyser-method. *Cytologia* 19: 336–357.

Kihara H. 1959. Fertility and morphological variation in the substitution backcrosses of the hybrid *Triticum vulgare × Aegilops caudate. Proceeding X International Congress of Genetics* 1: 142–171.

Kimber G, Tsunewaki K. 1988. Genome symbols and plasma types in the wheat group. In: Miller TE, Koebner RMD, eds. *Proceedings VII international wheat genetics symposium, Cambridge, UK.* Cambridge, UK: Cambridge University Press, 1209–1210.

Lilienfeld F. 1951. Genome-analysis in *Triticum* and *Aegilops.* X. Concluding review. *Cytologia* 16: 101–123.

Marcussen T, Sandve SR, Heier L, Spannagl M, Pfeifer M, The International Wheat Genome Sequencing Consortium, Jakobsen KS, Wulff BBH, Steuernagel B, Mayer KFX *et al.* 2014. Ancient hybridization among the ancestral genomes of bread wheat. *Science* 345: 1250092.

Matsuoka Y. 2011. Evolution of polyploid *Triticum* wheats under cultivation: the role of domestication, natural hybridization and allopolyploid speciation in their diversification. *Plant and Cell Physiology* 52: 750–764.

Meimberg H, Rice KJ, Milan NF, Njoku CC, McKay JK. 2009. Multiple origins promote the ecological amplitude of allopolyploid *Aegilops* (Poaceae). *American Journal of Botany* 96: 1262–1273.

Middleton CP, Senerchia N, Stein N, Akhunov ED, Keller B, Wicker T, Kilian B. 2014. Sequencing of chloroplast genomes from wheat, barley, rye and their

relatives provides a detailed insight into the evolution of the Triticeae Tribe. *PLoS ONE* **9**: e85761.

Pazy B, Zohary D. 1965. The process of introgression between *Aegilops* polyploids: natural hybridization between *A. variabilis*, *A. ovata* and *A. biuncialis*. *Evolution* **19**: 385–394.

Petersen G, Seberg O, Yde M, Berthelsen K. 2006. Phylogenetic relationships of *Triticum* and *Aegilops* and evidence for the origin of the A, B, and D genomes of common wheat (*Triticum aestivum*). *Molecular Phylogenetics and Evolution* **39**: 70–82.

Soltis DE, Segovia-Salcedo MC, Jordon-Thaden I, Majure L, Miles NM, Mavrodiev EV, Mei W, Cortez MB, Soltis PS, Gitzendanner MA. 2014. Are polyploids really evolutionary dead-ends (again)? A critical reappraisal of Mayrose *et al.* (2011). *New Phytologist* **202**: 1105–1117.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* **30**: 2725–2729.

Terachi T, Ogihara Y, Tsunewaki K. 1987. The molecular basis of genetic diversity among cytoplasms of *Triticum* and *Aegilops*. VI. Complete nucleotide sequences of the *rbcL* genes encoding H- and L-type Rubisco large subunits in common wheat and *Ae. crassa* 4×. *Japanese Journal of Genetics* **62**: 375–387.

Tsunewaki K. 2009. Plasmon analysis in the *Triticum*–*Aegilops* complex. *Breeding Science* **59**: 455–470.

Waines JG, Johnson BL. 1972. Genetic differences between *Aegilops longissima*, *A. sharonensis*, and *A. bicornis*. *Canadian Journal of Genetics and Cytology* **14**: 411–415.

Wang GZ, Miyashita NT, Tsunewaki K. 1997. Plasmon analyses of *Triticum* (wheat) and *Aegilops*: PCR-single strand conformational polymorphism (PCR-SSCP) analyses of organellar DNAs. *Proceedings of the National Academy of Sciences, USA* **94**: 14570–14577.

Yakimowski SB, Rieseberg LH. 2014. The role of homoploid hybridization in evolution: a century of studies synthesizing genetics and ecology. *American Journal of Botany* **101**: 1247–1258.

Yamane K, Kawahara T. 2005. Intra- and interspecific phylogenetic relationships among diploid *Triticum*–*Aegilops* species based on base pair substitutions, indles and microsatellites in chloroplast non-coding sequences. *American Journal of Botany* **92**: 1887–1898.

Zohary D, Feldman M. 1962. Hybridization between amphidiploids and the evolution of polyploids in the wheat (*Aegilops*–*Triticum*) group. *Evolution* **16**: 44–61.

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 Topologies of the 275 nuclear genes used in Marcussen *et al.* (2014), outgroup rooted (outgroups not shown).

Please note: Wiley Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

About *New Phytologist*

- *New Phytologist* is an electronic (online-only) journal owned by the New Phytologist Trust, a **not-for-profit organization** dedicated to the promotion of plant science, facilitating projects from symposia to free access for our Tansley reviews.

- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as ready' via *Early View* – our average time to decision is <27 days. There are **no page or colour charges** and a PDF version will be provided for each article.

- The journal is available online at Wiley Online Library. Visit **www.newphytologist.com** to search the articles and register for table of contents email alerts.

- If you have any questions, do get in touch with Central Office (np-centraloffice@lancaster.ac.uk) or, if it is more convenient, our USA Office (np-usaoffice@lancaster.ac.uk)

- For submission instructions, subscription and all the latest information visit **www.newphytologist.com**