Wheat speciation and adaptation: perspectives from reticulate evolution

Authors

Xuebo Zhao[1,2], Xiangdong Fu[1,2], Changbin Yin[1,*], Fei Lu[1,2,3,*]

Affiliations

[1]State Key Laboratory of Plant Cell and Chromosome Engineering, Institute of Genetics and Developmental Biology, Innovative Academy of Seed Design, Chinese Academy of Sciences, Beijing, China.

[2]University of Chinese Academy of Sciences, Beijing, China.

[3]CAS-JIC Centre of Excellence for Plant and Microbial Science (CEPAMS), Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China.

Correspondence
flu@genetics.ac.cn (F.L.); yinchangbin@genetics.ac.cn (C.Y.)

Abstract

Reticulate evolution through the interchanging of genetic components across organisms can impact significantly on the fitness and adaptation of species. Bread wheat (*Triticum aestivum* subsp. *aestivum*) is one of the most important crops in the world. Allopolyploid speciation, frequent hybridization, extensive introgression, and occasional horizontal gene transfer (HGT) have been shaping a typical paradigm of reticulate evolution in bread wheat and its wild relatives, which is likely to have a substantial influence on phenotypic traits and environmental adaptability of bread wheat. In this review, we outlined the evolutionary history of bread wheat and its wild relatives with a highlight on the interspecific hybridization events, demonstrating the reticulate relationship between species/subspecies in the genera *Triticum* and *Aegilops*. Furthermore, we discussed the genetic mechanisms and evolutionary significance underlying the introgression of bread wheat and its wild relatives. An in-depth understanding of the evolutionary process of *Triticum* species should be beneficial to future genetic study and breeding of bread wheat.

Keywords: wheat; reticulate evolution; introgression; speciation; hybridization; adaptation; breeding

Introduction

The 'tree of life' is a classic paradigm representing the evolution of life and illustrating the relationship of species[1]. While being practical and useful, the tree-like model can be incomplete because fusion, transfer, and exchange of DNA between species can form a net-like evolutionary relationship, shaping the paradigm of reticulate evolution[2]. Recently, a growing body of research shows that reticulate evolution is common in nature – hybridization, gene flow, and polyploidization were identified in multiple species of humans[3–6], animals[7–15], and plants[16–20]. The drastic turnover of genetic components brings forth rapid changes of genetic diversity, diversification, and speciation, establishing a new fitness landscape of species[7,14]. For crops, gene flow through reticulate evolution can be a valuable resource to investigate genetic mechanisms and candidate genes of crops adapting to various environments[21].

Wheat cultivation began from the Neolithic Age. Over the past few thousand years, wheat was of extreme importance as a staple food that facilitated human civilization development[22]. In the present-day, wheat has become the second most-produced cereal crop and the most widely grown crop in the world, contributing ~20% of the calories and proteins to the human diet[23]. However, climate models predict an acceleration of environmental extremes in near future, which are probably the prelude to a subsequent shortage of agricultural supply worldwide[24]. Moreover, the climate resilience of wheat was found declining in most European countries during the last 10 years, probably due to the reduced diversity in the genetic pool of cultivars[25,26]. Improving the ability of crops to cope with both biotic and abiotic stresses will likely play an important role in the adaptation of agriculture to climate change in the coming decades[27,28]. To address the agricultural challenges of this century, there is a pressing demand to understand the adaptive evolution of wheat and its wild relatives, and better utilize the gene pool of wheat.

Reticulate evolution of wheats (*Triticum* species) has long been recognized. The speciation of wheats is a classic example of reticulate evolution through allopolyploidization, shaping tetraploid and hexaploid species in the genus *Triticum*[29–31]. Two economically important wheats are bread wheat (*Triticum aestivum* subsp. *aestivum*, AABBDD, $2n = 6x = 42$) and durum wheat (*T. turgidum* subsp. *durum*, AABB, $2n = 4x = 28$), comprising 95% and 5% of the global wheat production, respectively[32,33]. Research of wheat evolution has focused primarily on two successive rounds of polyploidization events of bread wheat[34,35]. Homoploid hybridization, introgression, and their biological significance received considerably less attention. Recent studies showed that bread wheat received introgression from their wild relatives, facilitating its adaptation to new environments[30,31] and resistance to biotic stress[36]. The benefits of wild species as readily available resources of adaptive alleles are getting recognized[37]. Here, we review the process of reticulate evolution of *Triticum* species, then discuss the biological significance of reticulate evolution in bread wheat and its wild relatives, with a highlight on the impact of introgression upon global adaptation of wheat.

Phylogenetic tree of *Triticum* species

The *Triticum*-*Aegilops* alliance, which contains the direct genetic donors of bread wheat, belongs to the tribe Triticeae. The tribe Triticeae emerging about 25 million years ago (Mya), was split into two subtribes, Hordeineae and Triticineae, about 15 million years ago[38,39]. Triticineae is an economically important subtribe, giving rise to the domesticated cereals of wheat, rye, and several important forage grass[33] (Fig. 1A). Bread wheat and its closely related species (Table 1) include 9 species and 29 subspecies, growing mostly in temperate zones, principally in the northern hemisphere[40,41]. These subspecies are annuals and have solitary spikelet, with their original habitats in the eastern Mediterranean, and the cradle of agriculture–Fertile Crescent[42,43].

Table 1 The nomenclature and main traits of bread wheat and their wild relatives in *Triticum* and *Aegilops*

Species	Subspecies	Genome constitution	Pollination mode*	Reference genome	Breeding state	Brittle rachis	Free-threshing	Names in this review (common name)
Ae. speltoides Tausch	*Ae. speltoides* subsp. *speltoides*	SS	C	No	Wild	Yes	No	Speltoides
	Ae. speltoides subsp. *ligustica* (Savign.) Zhuk.	SS	C	No	Wild	Yes	No	Ligustica
Ae. tauschii	*Ae. tauschii* Coss. subsp. *strangulata*	DD	S	Yes	Wild	Yes	No	Strangulata
	Ae. tauschii Coss. subsp. *tauschii*	DD	S	No	Wild	Yes	No	Tauschii
T. monococcum	*T. monococcum* L. subsp. *aegilopoides* (Link) Thell.	AA	S	No	Wild	Yes	No	Wild einkorn
	T. monococcum L. subsp. *monococcum*	AA	S	No	Domesticated	No	No	Domesticated einkorn
T. sinskajae	*T. sinskajae* A.Filat. & Kurk. subsp. *sinskajae*	AA	S	No	Domesticated	No	No	Sinskaya
T. urartu	*T. urartu* Tumanian ex Gandilyan	AA	S	Yes	Wild	Yes	No	Urartu
T. turgidum	*T. turgidum* L. subsp. *dicoccoides* (Korn. ex Asch. & Graebn.) Thell.	AABB	S	Yes	Wild	No	No	Wild emmer
	T. turgidum L. subsp. *dicoccon* (Schrank) Thell.	AABB	S	No	Domesticated	No	No	Domesticated emmer
	T. karamyschevii NEVSKI var. *karamyschevii*	AABB	S	No	Domesticated	No	No	Georgian wheat
	T. ispahanicum Heslot	AABB	S	No	Domesticated	No	No	Ispahanicum
	T. turgidum L. subsp. *turgidum*	AABB	S	No	Domesticated	No	Yes	Rivet wheat
	T. turgidum L. subsp. *polonicum* (L.) Thell.	AABB	S	No	Domesticated	No	Yes	Polish wheat
	T. turgidum L. subsp. *carthlicum* (Nevski) A. Love & D. Love	AABB	S	No	Domesticated	No	Yes	Persian wheat
	T. turgidum L. subsp. *turanicum* (Jakubz.) A. Love & D. Love	AABB	S	No	Domesticated	No	Yes	Khorasan wheat
	T. turgidum L. subsp. *durum* (Desf.) Husn.	AABB	S	Yes	Domesticated	No	Yes	Durum
T. timopheevii	*T. timopheevii* subsp. *armeniacum (Jakubz.) van Slageren*	AAGG	S	No	Wild	Yes	No	Wild timopheevii
	T. timopheevii subsp. *timopheevii*	AAGG	S	No	Domesticated	No	No	Domesticated timopheevii
T. aestivum	*T. aestivum* L. subsp. *spelta* (L.) Thell.	AABBDD	S	Yes	Wild/Domesticated	No	No	Spelt
	T. aestivum L. subsp. *macha* (Dekapr. & Menabde) Mackey	AABBDD	S	No	Wild/Domesticated	No	No	Macha
	T. aestivum L. subsp. *compactum* (Host) Mackey	AABBDD	S	No	Domesticated	No	Yes	Club wheat
	T. aestivum L. subsp. *sphaerococcum* (Percival) MacKey	AABBDD	S	No	Domesticated	No	Yes	Indian dwarf wheat
	T. aestivum subsp. *yunanense*	AABBDD	S	No	Domesticated	No	No	Yunan wheat
	T. petropavlovskyi Udachin & Migush.	AABBDD	S	No	Domesticated	No	Yes	Xinjiang wheat
	T. aestivum subsp. *tibetanum*	AABBDD	S	No	Wild/Domesticated	Yes	No	Tibetan semi-wild
	T. vavilovii (Tumanian) Jakubz. var. vavilovii	AABBDD	S	No	Wild/Domesticated	No	Yes	Vavilovii
	T. aestivum L. subsp. *aestivum*	AABBDD	S	Yes	Domesticated	No	Yes	Bread wheat
T. zhukovskyi	*T. zhukovskyi Menabde & Ericz.*	AAAAGG	S	No	Wild	No	No	Zhukovskyi

Source: Wheat Genetics Resource Center (https://www.k-state.edu/wgrc/genetic_resources/triticum_accessions/index.html).
* C cross-pollination, S self-pollination

During the speciation of bread wheat, the first hybridization occurred ~0.8 Mya between close relatives (BB) *of Ae. speltoides* and *T. urartu* (AA), giving rise to the allotetraploid wild emmer wheat (*T. turgidum* subsp. *dicoccoides*, AABB) by polyploidization. At the beginning of agriculture in the Fertile Crescent about 10,000 years ago, wild emmer was domesticated and their cultivated forms appeared[44]. Domesticated emmer wheat (*T. turgidum* subsp. *dicoccon*) was the product of this initial process with the non-brittle rachis trait. Then, domesticated emmer went through the second stage of domestication and free-threshing tetraploids emerged with naked grain phenotype[34]. As of today's economically important wheat, durum is a subspecies of free-threshing tetraploids[32]. Recent studies showed that bread wheat originated through allopolyploidization from the second hybridization event, between free-threshing tetraploids and *Ae. tauschii* (DD)[31]. To date, no wild forms of *T. aestivum* have been identified, indicating that these hexaploid species are innate cultivars[45]. Bread wheat landraces are clearly divided into Asian and European groups. Almost all bread wheat varieties are derived from the intensive use of European wheat, suggesting that recent dispersal and selection have resulted in a high imbalance in modern germplasm compared to the ancestors of local varieties[31,46] (Fig. 1B).

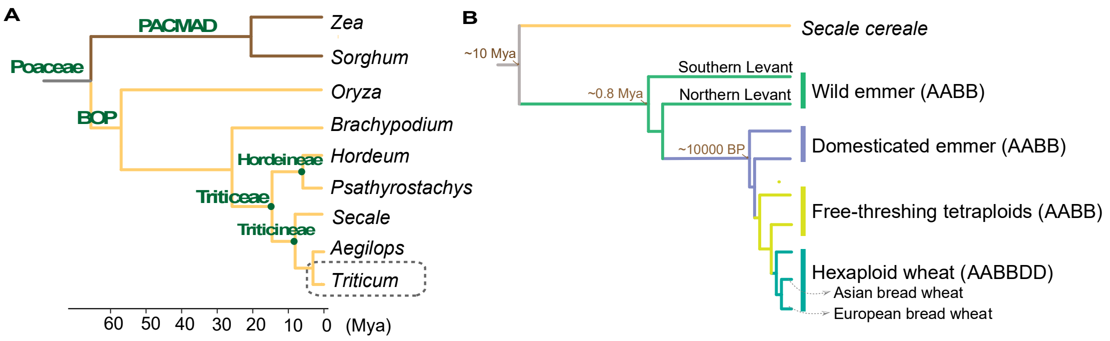

Fig.1 Phylogenetic tree showing the evolutionary relationship of *Triticum/Aegilops* species. A. Phylogeny of Triticeae and its position in the family Poaceae. **B.** A pruned tree of *Triticum* species illustrating the process of domestication and speciation of hexaploid wheat.

Reticulated model completes the phylogenetic relationship of wheats

Natural hybridization is defined as the intercross of individuals from two different species or populations of the same species, while introgression or introgressive hybridization describes the incorporation of alleles from one population into another via hybridization followed by continuous backcrossing[47]. New subspecies formation in the *Triticum-Aegilops* alliance occurs frequently by polyploidization as well as intra- and inter- introgression via hybridization, leading to a rapid burst of diversification[48–50]. For example, the speciation of bread wheat is a result of two successive rounds polyploidization (Fig. 2A). There are several reasons that reticulated evolution can happen in *Triticum*/*Aegilops* species. First, in the middle and the end of Pleistocene, diploid donors and tetraploid donors formed the allopolyploid species of *Triticum* by genome duplication via hybridization, possibly due to the global cooling during primary ice ages[51,52]. Second, being relatively young species, the *Triticum-Aegilops* alliance showed an exceptional hybridization capacity involving most of its members, most of which have not yet reached complete reproductive isolation at the species level[38,50]. Third, wheat with different ploidy levels and domestication statuses were grown in the same field. Hybridization and introgression provided a wide range of phenotypes and made possible a selection of desirable genotypes to the farmers. For thousands of years, farmers have been growing mixtures of different genotypes and even different cytotypes[53,54]. Major hybridization events are as follows.

Hybridization events in the diploid *Triticum/Aegilops* species

Cultivated wheats are derived from an intricate history of four basic genome types, A, B, G, and D, presented in both diploid and polyploid species. The diploid *Triticum* and *Aegilops* species started to diverge from one another during the middle of Pliocene, about 2 to 4 Mya[38]. The genus *Triticum* consists of the A-genome diploids *T. urartu* ($2n = 2x = 14$, A^uA^u) and *T. monococcum* subsp. *aegilopoides* ($2n = 2x = 14$, AA). Similar to the diploid *Triticum* species, the genome of diploid *Aegilops* species, including *Ae. tauschii* Coss. ($2n = 2x = 14$, DD) and *Ae. speltoides* Tausch ($2n = 2x = 14$, SS), evolved from the 7 ancestral chromosomes of *Triticum-Aegilops* alliance[55,56]. The B genome and G genome in *Aegilops* might be referred to as S genome in other treatments. Nevertheless, the B, G, and S genomes, though related, are quite distinct[40].

Hybridization events raised long-standing controversies in the classification of diploid *Triticum/Aegilops* species. Marcussen et al.[55] and Baidouri et al.[57] proposed that the D genome originated from an ancient hybridization event between the A and B lineages about 2.1~2.9 Mya. However, using transcriptome data phylogenomic analysis, Sylvain et al.[58] suggested that another wild species *Ae. Mutica* may be closer to D genome's ancestor than B genome and B genome has continuous introgression into D genome[59] (Fig. 2B, label a). In addition, by comparing with the *Triticum urartu* (AA) and *Aegilops tauschii* (DD), whether the B genome in bread wheat originated from *Ae. speltoides* (SS) is also in intense debate. The findings from earlier studies suggested two hypotheses. First, the progenitor of the B genome is a unique and ancient *Aegilops* species that remains as an unknown ancestor closely related to *Ae. speltoides*[60]. Second, the progenitor of the B genome is *Ae. speltoides*, but the genome of extant *Ae. speltoides* comprises segments of several species in *Aegilops* introgressed into *Ae. speltoides*[61]. Due to the self-incompatible nature of *Ae. speltoides*, ongoing genome recombination in its natural population makes the origin of the B genome more bewildering.

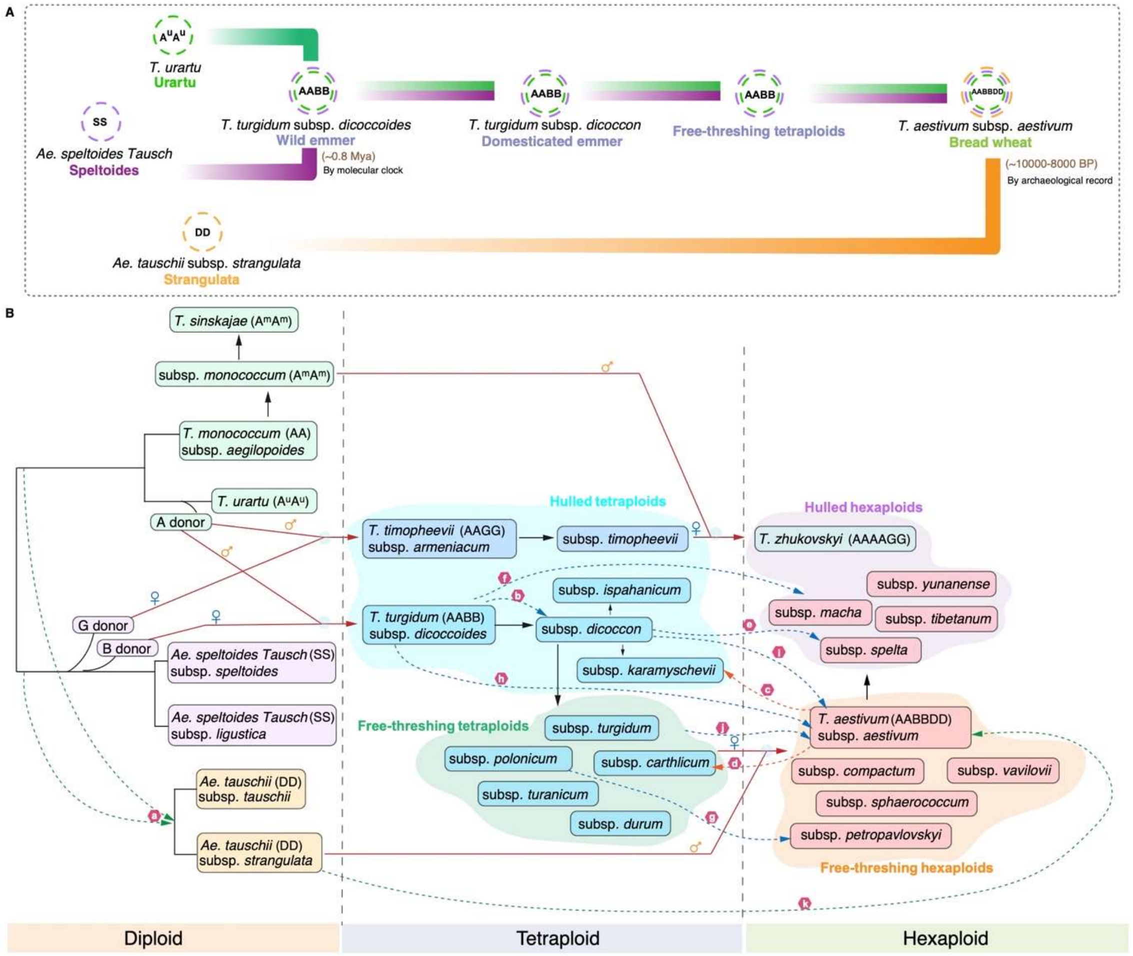

Fig.2 The reticulate evolutionary relationship between bread wheat and its progenitors in *Triticum/Aegilops* species. A. Schematic diagram of the speciation process of bread wheat through two successive rounds polyploidization. B. Model of reticulated evolution involving *Triticum*/*Aegilops* species. The diploid, tetraploid, and hexaploid species are separated by gray dotted line. The red arrows show occurrences of allopolyploid hybridization events, which are labeled from 1 to 4. The black arrows show domestication events. The green, blue and orange dotted line show the homoploid hybridization events from diploids, tetraploids and hexaploids, which are labeled from a to k. The areas of blue, green, purple, and orange indicate hulled tetraploids, free-threshing tetraploids, hulled hexaploids, and free-threshing hexaploids, respectively.

Speciation through polyploidization in *Triticum*

Allopolyploidization is one of the driving forces creating new species[62,63]. Allotetraploid wheat *T. turgidum* (genome AABB) originated from hybridization events involving two diploid progenitors classified in the genera *Aegilops* and *Triticum* (Fig. 2B, label 1). Genome B, as described above, which was derived from *Ae. speltoides* or from a closely related species to *Ae. speltoides*, underwent changes of the ploidy level. Genome A, which has been modified relatively little, was derived from a sister species of *T. urartu.* Another allotetraploid wheat *T. timopheevii* (genome AAGG) contains a pair of A genomes from *T. urartu* and a pair of G genomes[55,64] (Fig. 2B, label 2). *T. turgidum* and *T. timopheevii* were domesticated, and their cultivated forms appeared during the early stage of agriculture in the Fertile Crescent about

10,000 years ago. These newly formed tetraploid wheats then went through extensive diversification due to human selection. The evolution of the *T. turgidum* as crops was initiated when wild emmer was domesticated and formed domesticated emmer wheat (*T. turgidum* subsp. *dicoccon*) and its varieties, *T. turgidum* subsp. *karamyschevii* and *T. turgidum* subsp. *ispahanicum*[65,66]. As the taxa representing the second stage of domestication, free-threshing tetraploids were derived from domesticated emmer, such as the extinct tetraploid *T. turgidum* subsp. *parvicoccum*, appearing in the archaeological record shortly after domesticated emmer[67]. As the most well-known free-threshing tetraploids, durum wheat is used for macaroni and semolina products[32]. Most other tetraploid wheat subspecies are free-threshing and probably arose relatively recently, including *T. turgidum* subsp. *turgidum, T. turgidum* subsp. *turanicum, T. turgidum* subsp. *polonicum, T. turgidum* subsp. *carthlicum*, all of which are quite similar to durum wheat and differ by only a few traits[40]. Likewise, the diploid *T. monococcum* also was domesticated in the same area from wild einkorn (*T. monococcum* subsp. *aegilopoides*) to domesticated einkorn (*T. monococcum* subsp. *monococcum*)[22,43]. However, they did not contribute to the speciation of tetraploid species of *Triticum* through polyploidization.

Speciation through hexaploidization also happens in the genus *Triticum*. *T. aestivum* (genome AABBDD) is thought to have arisen through hybridization between *T. turgidum* and *Aegilops tauschii* (Fig. 2B, label 3), while *T. zhukovskyi* (genome AAAAGG) originated through hybridization between *T. timopheevii* and domesticated einkorn *T. monococcum* (Fig. 2B, label 4)[68–70]. *Ae. tauschii* subsp. *strangulata* is the accepted ancestor of the D subgenome of bread wheat[71], but the tetraploid donor was not clear until a recent study showed that AB subgenomes of bread wheat originated from free-threshing tetraploids[31]. To date, no wild forms are identified for either *T. aestivum* or *T. zhukovskyi*., indicating that these hexaploid species are innate cultivars. Other hexaploid wheats include 3 subspecies based on morphological traits–*T. aestivum* subsp. *compactum*, *T. aestivum* subsp. *macha*, and *T. aestivum* subsp. *spelta*, and 5 subspecies named according to geographical location–*T. aestivum* subsp. *sphaerococcum*, *T. aestivum* subsp. *Yunanense*, *T. aestivum* subsp. *petropavlovskyi*, *T. aestivum* subsp. *tibetanum,* and *T. aestivum* subsp. *vavilovii*[40,41].

Speciation through homoploid hybridization in *Triticum*

Hybridization creates new species not only through polyploidization but also from homoploid hybridization[72,73], in which the chromosome number is not changed in the offspring. In tetraploids, gene flow from wild populations in southern Levant contributed to the population structure of domesticated emmer, but domesticated emmer actually originated from northern Levant[65,74] (Fig. 2B, label b). Moreover, *T. turgidum* subsp. *karamyschevii* are considered to be a segregant from a hybrid between wild emmer and bread wheat[75] (Fig. 2B, label c), and Persian wheat (*T. turgidum* subsp. *carthlicum*) may receive the introgression from hexaploid wheats[76] (Fig. 2B, label d).

With regards to the origin of hexaploid wheats, a once-popular idea was that spelt wheat was the primitive form of hexaploid wheats because the initial hexaploid wheat was expected to have hulled seed inherited from *Ae. tauschii*, and therefore would be morphologically similar to spelt[77]. However, HMW glutenin analysis showed a result of hybridization between *T. aestivum* and domesticated emmer as the origin of European spelt[78–80] (Fig. 2B, label e). Macha wheat (*T. aestivum* subsp. *macha*) provides another example of bread wheat diversification through homoploid introgression[75] (Fig. 2B, label f), and the genome of Xinjiang wheat (*T.* subsp. *petropavlovskyi*) has many introgressed segments from free-threshing tetraploids[81] (Fig.

2B, label g). In bread wheat, introgressed segments were found from their diploid and tetraploid progenitors, accounting for a substantial portion of the bread wheat genome[74]. The AB subgenomes of bread wheat received large amounts of introgression from multiple wild tetraploid populations, such as wild emmer (Fig. 2B, label h), domesticated emmer (Fig. 2B, label i), and free-threshing tetraploids (Fig. 2B, label j). However, relatively few *Ae. tauschii* introgression (Fig. 2B, label k) was introduced into the D subgenome of bread wheat. To date, there is no evidence showing that *T. urartu* contributed gene flow to bread wheat[31,82]. This is possibly due to the low survival rate of unreduced gametes and hybrid offspring, as well as the narrow geographical distribution of ancestral diploid populations[41,83]. An introgressed segment of 150Mb from wild emmer to bread wheat on the 4A chromosome was identified[30]. The largest introgression segment was from free-threshing tetraploids, which was on chromosome 2A of the bread wheat genome, covering ~50% of the whole chromosome[31].

Methods of introgression detection

Compared with polyploidization with altered chromosome number, introgression segments are more difficult to detect at the homoploid level, because ancestral polymorphisms are shared across many descendant species, which is referred to as incomplete lineage sorting (ILS) and it is challenging to be distinguished from introgression[21,84]. Furthermore, ancient introgression will dilute over generations[85]. Whole-genome data maximize the power to detect introgression, even if the introgressed segments are rare in the genome[17,86–88]. Many useful statistics have been developed to detect introgression.

At the whole-genome level, STRUCTURE[89], BEAST[90] and phyloNet[91] could be used to determine whether hybridization events have occurred. Besides, approximate bayesian computation model such as DIYABC[92] estimate introgression rate as a component of broader demographic history. At the locus level, three categories of methods have been developed based on the extent of differentiation, phylogenetic relationships, and demographic history. First, the $\mathrm{RND}_{\mathrm{min}}$[93] are most used based on F_{st} and d_{XY} statistics. These differentiation-based approaches can be calculated site by site, dense genome-wide data are not required. However, they do not polarize the introgression events. Second, the f_3, f_4 statistics[94], D-statistic or ABBA-BABA test[95], HyDe[96] and five taxon D statistics[97] are based on phylogenetic relationships, which is capable of detecting the location and size of introgression, as well as the direction of gene flow, but they have limited ability to distinguish between introgression and ILS. Third, methods such as F statistics[98] and S∗ statistic[99], coalescent simulations of specific demographic scenarios can obtain the direction of gene flow, in which the information of phylogeny is not required. This group of methods can distinguish introgression and ILS based on the assumption that introgression haplotypes should be larger than ILS haplotypes because recombination can break down haplotypes over generations[100]. However, the disadvantage of these methods is that the demographic history of species is relatively complex, leading to inaccurate estimation of ancestral state[87]. Collectively, introgression detection methods have been successfully applied to several systems[5,9,13,21,84].

Role of Introgression in the Adaptation of Wheat

Several types of hybridization events have been detected in wheats suggesting that the reticulate evolution pattern largely complete the hierarchical structure of relationships in *Triticum-Aegilops* Alliance. Hybridization and introgression had an important role in the speciation and diversification of wheats, contributing to the flourishing of the genus *Triticum* and the unclear boundaries between the subspecies. Notably, domesticated tetraploid and hexaploid wheats exhibit a wider range of genetic plasticity than their progenitors and could adapt themselves to various environments[29]. By integrating subgenomes previously existing in different environments, polyploid genome brings about a broader range of adaptability of allopolyploid wheat[52,101,102]. However, the polyploidization events induced severe diversity reduction in the ancestral population of allopolyploid wheat, which was likely hindering its range expansion[28,103]. Therefore, an in-depth and timely discussion about the role of reticulate evolution in the adaptation of wheats is valuable. As the biological significance of homoploid hybridization and introgression has not been thoroughly studied in wheat, we will use examples from a broad range of species to discuss the role of introgression in the adaptation of wheat in the following section.

Recent studies showed that the tremendous diversity loss of bread wheat, during the bottleneck events of polyploidization and domestication, was mainly compensated by extensive introgression from wild populations[30,31,74,82] (Fig. 3A). These introgression events captured a considerable proportion of phenotypic variability of both diploid and tetraploid progenitors, granting an incredible potential of bread wheat adapting to human selection in diverse environments. However, pros and cons of introgression may both exist and impact on the adaptation of wheat. To make the most effective introduction of desirable alien alleles into cultivars, it is necessary to understand the genetic mechanism and evolutionary significance of introgression, especially the benefit and cost of introgressed alleles to the recipient population (Fig. 3B).

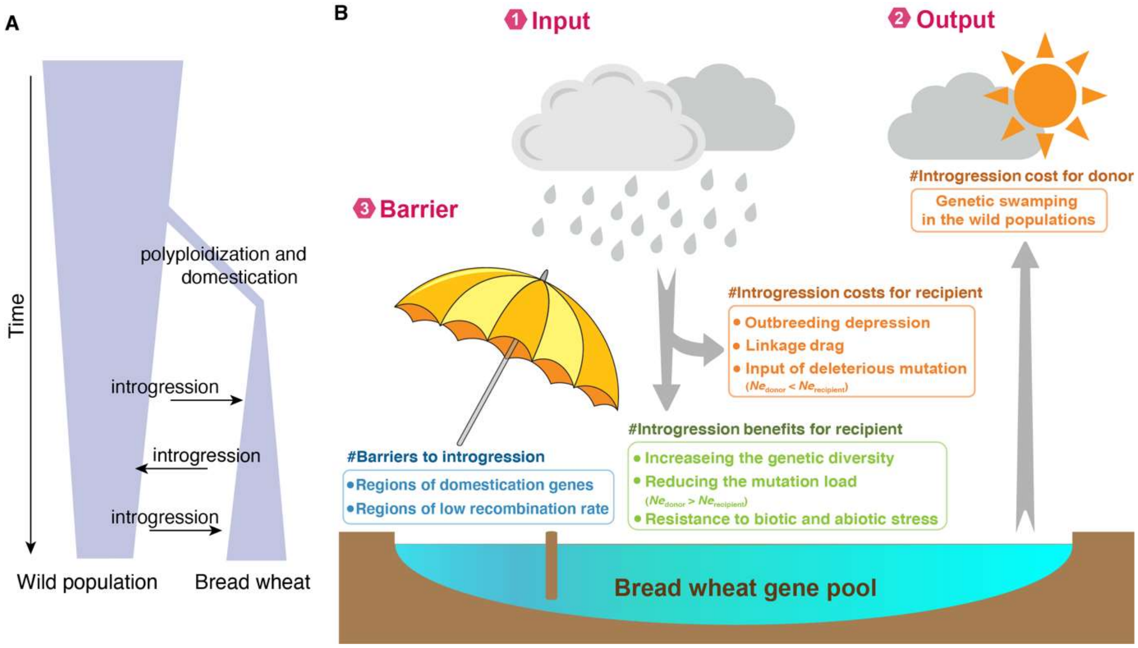

Fig.3 The process and genetic mechanisms of introgression between bread wheat and wild populations. A. Simplified demographic model of wheat populations. The size of the branches reflects effective population sizes (*Ne*). Arrows indicates introgression. **B.** Genetic mechanisms and effect of introgression in wheats.

The benefits of introgression

Compared with domesticated crops, wild populations are characterized by high diversity, large effective population size, and higher fitness in the natural environment[28,104]. Introgression from wild relatives contributes to the divergent adaptation of allopolyploid wheat by substantially increasing its genetic diversity, reducing the deleterious mutation, and expand tolerance to extreme environments[14,17,105]. When these foreign introgression alleles increase the fitness of the recipient pool is called “adaptive introgression”[21] (Fig. 3B label 1). The adaptive introgression can happen in several ways, which are listed as follows.

Increasing the genetic diversity. Introgression increases the genetic diversity by up to two to three orders of magnitude more than mutations[106,107]. In many crops, such as rice[108] and soybean[18], introgression from wild populations into cultivated species could increase the diversity of cultivated populations substantially. Many researches have demonstrated that the frequent introgression contributed to the restoration of genetic diversity in bread wheat[30,31,74,82]. Tetraploid introgression to hexaploid was more frequent and relatively recent, probably due to the higher survival rate of the offspring after homologous hybridization. The tetraploid donors include wild emmer wheat, domesticated emmer wheat, and free-threshing tetraploid wheats, which contributes more than 20% of the bread wheat genome in most cultivars. By contrast, introgression from diploid donor (genome DD) into bread wheat was little, accounting for less than 2% of the bread wheat genome. The imbalance of the introgression from tetraploid and diploid donors is the reason for the asymmetric distribution of nucleotide diversity on the ABD subgenomes with $\pi_B > \pi_A \gg \pi_D$[31,74]. More than 60% nucleotide diversity of wild emmer was captured by composite gene flow from tetraploids in bread wheat genome. On the contrary, bread wheat preserved only 14% of the nucleotide diversity of its D subgenome donor, strangulata, which led to the extremely low level of nucleotide diversity in the D subgenome of bread wheat[31]. Thus, we speculate that the subtle nucleotide diversity differences in A and B subgenomes may also resulted from the different level of gene flow in tetraploid wheats from their ancestral populations *T. urartu* and *Ae. speltoides*. In addition to the introgression from the genera *Triticum* and *Aegilops*, segments introgressed from barley and rye into bread wheat also contributed to the increase of diversity[37,109,110].

Reducing the mutation load. Mutation load refers to the accumulation of deleterious alleles across the genome, which reduces the fitness of species[111]. The introgression regions from wild populations can result in a reduction in deleterious mutation burden compared with other regions. Since a larger effective population size of wild relatives has the capability to remove deleterious alleles effectively, introgressed alleles from wild relatives can offset the bottleneck effect and increase the selective efficacy[112]. By quantifying deleterious mutation load using additive or recessive models, the maize population showed a strong negative correlation between mutation load and introgression[113]. In wheat, introgression of wild emmer wheat showed reduced mutation load in hexaploid landraces and cultivars[30].

Resistance to biotic and abiotic stress. The introgressed alleles associated with local adaptation could be the key to evolutionary success to local conditions and expansion around the world[82]. In potato, introgression is more likely to be highly expressed and stress-induced. The introgressed region contains loci associated with disease resistance, drought tolerance, and heat resistance, suggesting that introgression confers key adaptations for survival and may promote tolerance to new environmental pressures during range expansion[114]. Introgression regions from wild emmer into bread wheat explained a substantial portion of phenotypic variance for harvest weight, plant height, and flowering time traits, whereas the introgression regions from

Ae. tauschii are related to disease resistance generally, indicated different roles of the three subgenomes during selection and adaptation[30,82]. Given that biotic and abiotic stress resistances are complex traits, high-throughput phenotyping should help identify functional introgression to cope with a changing climate[115,116].

The costs of introgression

The benefits of introgression certainly increase the genetic diversity and enhance the plasticity of wheat adapting to human selection in different environments. However, every coin has two sides. We also have to be soberly aware that gene flow involving domesticated species is not universally beneficial. Genetic background, demographic history, and environmental factors of both donor and recipient populations can affect the efficacy of introgression breeding. Introgression from wild populations also has negative effects to the population of bread wheat. We will discuss the costs of introgression as follow (Fig. 3B, label 1 and 2).

Outbreeding depression. Genetic differences between the parental species often result in a fitness drop in the hybrids, commonly deemed as “outbreeding depression”[117–119]. An extreme case of outbreeding depression is hybrid incompatibility, where the hybrid offspring become inviable or sterile[120]. Probably to ensure the gene dosage balance and mitigate the outbreeding depression, the newly formed allopolyploids usually undergo instantaneous elimination and rearrangements of DNA sequences during the first a few generations[103,117]. These rearrangements, including sequence deletion, transposons activation or deactivation, as well as pseudogenization, subfunctionalization, and neofunctionalization of genes, are extensive and rapid[121–123]. While wide hybridization has been used intensively in wheat improvement, these wild hybrids often show reduced grain yield and quality[37]. Triticale is a synthetic species by crossing wheat and rye with the purpose of combining the characteristics of high yield and high quality of wheat with disease resistance, cold resistance, and high lysine content of rye[124]. However, low seeding rate and shriveled grain are common traits in octoploid triticale, probably due to the outbreeding depression[125]. Dosage compensation in plants may provide a genetic mechanism making divergent genomes undergo strong interactions and domination in allopolyploid offspring[120,126]. In hexaploidy wheat, the Q locus is important to wheat domestication and maintains the phenotype of naked grain. The hyperfunctionalization of 5AQ produced the domestication phenotype in tetraploid wheats (AABB). After strangulata (D genome) joining in, 5Dq in bread wheat became subfunctionalized and showed a reduced gene expression, probably to offset the dosage perturbation caused by the addition of the D subgenome[127].

Linkage drag. Linkage drag refers to a phenomenon that undesirable genes are introduced into breeding pool with introgression segments of interests attributed to genetic linkage. Early farmers have made long-term and continuous effort in selecting agronomic traits in wheat in different areas[53]. The introgression from wild populations did increase genetic diversity of the recipient population, but due to the linkage drag, it often gives rise to weedy phenotypes and thus erode the fruits of thousands of years of domestication and improvement. For example, the famous wheat-rye 1BL.1RS translocation has been extensively used in wheat cultivars because of its remarkable effect of disease-resistance[128]. However, this introgressed segment also contains genes that reduce the processing and baking quality of wheat[129].

How are the existing adaptive alleles in bread wheat are maintained given the introgression and concomitant linkage drag occurs so commonly? The retention of adaptive alleles in crop populations is largely governed by selection. Studies have shown that introgression signals from wild relatives into domesticates are reduced around genes responsible for major domestication traits[30,113] (Fig. 2B label 2). The locus of wheat domestication gene *Q* is located in the distal region of 5A chromosome, which is characterized by high gene density and high recombination rate. However, the decrease in genetic diversity was found near the gene *Q* and no introgression from wild populations was detected in bread wheat, in accordance with the studies mentioned above[30,130]. Likewise, gene *Tg*, another locus controlling the free-threshing trait, locates also in a distal chromosome with frequent recombination and alien introgression, but shows reduced introgression only within ~10 Mb flanking region and strong negative selection[82,131] (Fig. 3B, label 3).

In addition to selection, recombination rate is an important factor determining the extent of introgression in the recipient population (Fig. 2B, label 3). In natural populations of butterfly, the rate the introgression can be predicted by recombination rate across different species [132]. From artificial intraspecies cross experiment in barley, recombination rate variation shaped the patterns of introgression genome-wide[133]. Moreover, recombination modifiers, especially structural variants, play a critical role in resisting introgression. In sunflower, reduced introgression around a large genomic inversion maintains the combination of adaptive alleles[116,134]. Similar phenomena have not been reported in wheat but the resistance to introgression due to recombination cold spot may also exist.

Genetic swamping in the wild populations. Genetic swamping mostly refers to a circumstance that gene flow from donor species leads to a gradual replacement of local recipient species by hybrids[135]. In crops, the gene flow between wild populations and cultivars has been found to be bidirectional and genetic swamping may happen in the wild species. The introduction of nontransgenic imidazolinone-resistant rice cultivars once helped boost rice yield by 50% in southern Brazil during 2003 to 2014. The herbicide resistant weedy rice started to replace the original weedy rice and evolve around 2006, thus more herbicide is required for weed control– the production cost was tripled at the same period and many farmers in the area went out business[136]. For maize, a highland wild relative of maize (*Zea mays* subsp. *mexicana*) contains about 10% of the genome introgressed from maize[17]. Gene flow from bread wheat to wild population has not been clear so far, but some of the wild relatives of bread wheat was proposed to be the offspring between bread wheat and wild populations[41]. Wild populations are extremely valuable as a natural germplasm bank for wheat improvement, the gene flow from cultivars may change the genetic make-up of wild populations or even put some endangered species at the risk of extinction[135,137]. The genetic effect of introgression from cultivars to the wild needs a thorough assessment.

Input of deleterious mutation. Introgressed alleles from donor to the recipient were sometimes deleterious, especially when the donor population has a very limited effective population size. This effect has been observed in the case of Neanderthal introgression into humans, in which the introgression segments was restricted primarily to nongenic regions because of the high mutation load found within Neanderthal genomes[138,139]. Hence, to increase efficacy of introgression breeding of wheat, it is crucial to investigate the demographic history of wild donor populations and evaluate their overall biodiversity and fitness.

The fate of introgressed segments

The fate of introgressed segments the recipient population largely depends on recombination rate and selective pressure[118]. In butterfly, introgression can be predicted by recombination rate across different populations[132]. In human, the extent of introgression is found to be associated with local recombination rate and introgression segments are enriched in highly conserved genomic regions[6,87]. Both lines of evidence demonstrate the key role of recombination in determining the amount of introgression. Moreover, introgression segments will decrease in size over time because of recombination. The time of introgression is negatively correlated with the size of introgression segment[140,141]. Once the gene flow was established, selection will act on the introgression segments and decide to a large extent whether these segments will be retained[30,87,142]. Simply speaking, if introgressed segments were beneficial, they are more likely to have increased allele frequency or even been fixed in the recipient population[73,87]. If the introgressed segment was deleterious, the segment would be purged by negative selection. Or if the segment is neutral, the allele frequency is likely to be fluctuating[87,139]. Although the genetic mechanisms determining the fate of introgressed alleles are derived from many different species, the same rules should be equally applying in wheat as well. For example, wide-hybridization between bread wheat and *Thinopyrum ponticum* was conducted to introduce rust resistance and boost yield in bread wheat. One of the introgression lines, Xiaoyan 6, was quite successful and wild grown in China in 1980s. At present-day, many elite wheat cultivars in China are derivatives of Xiaoyan 6, containing introgression segments of rust resistance[143,144].

The application of introgression in wheat breeding

Unlike mutations, gene flow makes it possible to introduce large numbers of new alleles into multiple loci simultaneously, which may be the source of functional alleles and even adaptation to polygenic traits[107]. The awareness of hybridization and introgression attracted researchers and breeders long ago, and the first successful artificial hybrid wheat was reported in the 19th century. In 1876, the Scottish botanist Stephen Wilson produced a wheat-rye hybrid, a milestone in crossing wheat with related species[145]. Beginning from the 1920s, improvements in in-vitro tissue culture methods, e.g. embryo rescue, allowed the success of previously unsuccessful hybrid combinations[146]. Since Sears coined the term "chromosome engineering" in 1972[147], approaches from molecular biology have been successively used to exploit the genetic diversity in different gene pools and develop new genetic resources derived from alien introgression[148,149].

Some disease-resistance and yield-related genes have been transferred from multiple levels of gene pools into wheat through homoeologous recombination, chromosome translocation and other approaches[37]. The primary gene pool, which is fully sexually compatible with bread wheat, has the same genomic composition comparing to bread wheat and contains three subgenomes of A, B and D. This gene pool consists of the subspecies in *T. aestivum*, such as spelt wheat, Indian dwarf wheat, and the varieties of bread wheat landrace[37,150]. The secondary gene pool comprises the majority of *Triticum* species and part of the *Aegilops* speices, which share homologous genomes with bread wheat[150,151]. For example, *Lr47*[152] and *Lr42*[151] are introgressed into bread wheat from *Ae. speltoides* and *Ae. tauschii*. Species in the tertiary gene pool belong to the Triticeae tribe, but do not share homologous genomes with bread wheat[37,54,153]. Rye and barley and representative speices in the tertiary gene pool. One of most notable examples of introgression breeding in wheat is the wheat-rye translocation 1BL.1RS, in which a genomic region in the short arm of wheat chromosome 1B is replaced by a genomic region in the short arm of rye chromosome 1R to enhance disease resistance of wheat[128,154]. In addition, *Sr52*[155], *Sr59*[156], *Pm56*[157], *Sr44*[158], and *β-glucan CslF*[159] are also introgressed to bread wheat from *H. villosa,* rye, *Th. Intermedium* and barley. The gene flow from hexaploid and tetraploid was increasing in wheat cultivars released before 1960, while alien introgression from Triticeae species had increasing contribution in the gene flow in cultivars released after 1960, which reflects the fact that breeders have spent tremendous effort in introducing align alleles into the wheat gene pool since the Green Revolution 60 years ago[153,160]. Despite these progresses, introgression breeding is still largely empirical, and more research is needed to achieve precise transfer of favorable alleles and efficient production of improved wheat without compromising other traits[46,153]. Recently, the high throughput genotyping technology increased the efficiency for the precise identifications of target segments from very large populations[30,31,74,160–164]. Besides, gene editing can realize the transfer of adaptive genes without cross experiment[165,166]. We anticipate that the availability of genomic tools and the full understanding of rich genetic resources will facilitate the future introgressions for wheat improvement.

Conclusion

The evolution of bread wheat and its wild relatives is a complex and ongoing process in dynamic environments. The most striking features of this evolution are the frequent hybridization events and extensive introgression that lead to diversification and adaptation [7,73]. This reticular pattern provides the genetic basis for blurred boundaries between species and a diverse gene pool of bread wheat. Thus, evaluating the wild introgression existing in the cultivated gene pool is important to understand wild diversity relevant for crop adaptation and inform new breeding strategies[21,84].

Review of previous studies has made it clear that many questions remain unsettled regarding the evolution of the wheat. For example, the evolutionary relationship between tetraploids and hexaploids at the subspecies level is unclear. Moreover, due to the short history of bread wheat and large amounts of introgression between bread wheat and its wild progenitors, the conventional taxonomic classification of subspecies in *Triticum* may overemphasizes their differences. Due to the same reasons, the progenitor-domestication-cultivar model applied in most crops may be oversimplified in wheat. There is probably something having been missed in terms of the origin, duration of domestication, and migration routes because of the ongoing gene flow.

In addition to hybridization history, questions about genetic mechanisms of introgression underneath wheat adaptation need to be addressed. The studies of gene flow in *Triticum* species at the whole-genome level just got off ground recently[31,37,74]. By combing with massive data from gene expression[167], epigenomics, and regulatory elements profiling[168–171], answers to these questions will soon be available. The development of chromosome engineering and genome editing makes it possible to introduce target alleles, multi-gene editing, and directed evolution of protein[166,172]. We believe an in-depth understanding of genetic rules and biological significance of reticulated evolution of *Triticum* species will help identify useful agronomic genes, facilitate precision breeding, and address challenges from climate change for wheat.

Acknowledgements

Due to limited space we apologize to participants whose excellent work not mentioned here. This work is supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA24020201 and XDA24040102) and the National Natural Science Foundation of China (31970631) to F.L., National Natural Science Foundation of China (31921005) to X.F.

Conflict of interest

All the authors have no conflict of interest statement.

Reference

1. Charles, D. Customers who viewed The Origin of Species : 150th Anniversary Edition also viewed The Origin of Species : 150th Anniversary Edition Mass Market Paperback –. (1859).
2. Linder, C. R., Moret, B. M. E. & Nakhleh, L. Network (reticulate) evolution: biology, models, and algorithms. *Pacific Symp.* 41 (2003).
3. Green, R. E. *et al.* A draft sequence of the neandertal genome. *Science* **328**, 710–722 (2010).
4. Reich, D. *et al.* Genetic history of an archaic hominin group from Denisova cave in Siberia. *Nature* **468**, 1053–1060 (2010).
5. Prüfer, K. *et al.* The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* **505**, 43–49 (2014).
6. Sankararaman, S. *et al.* The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* **507**, 354–357 (2014).
7. Edelman, N. B. *et al.* Genomic architecture and introgression shape a butterfly radiation. *Science* **599**, 594–599 (2019).
8. Verdugo, M. P. *et al.* Ancient cattle genomics, origins, and rapid turnover in the Fertile Crescent. *Science* **365**, 173–176 (2019).
9. Daly, K. G. *et al.* Ancient goat genomes reveal mosaic domestication in the Fertile Crescent. *Science* **361**, 24–27 (2018).
10. Frantz, L. A. F. *et al.* Erratum: Ancient pigs reveal a near-complete genomic turnover following their introduction to Europe (Proceedings of the National Academy of Sciences of the United States of America (2019) 116 (17231-17238) DOI: 10.1073/pnas.1901169116). *Proc. Natl. Acad. Sci. U. S. A.* **117**, 14610–14611 (2020).
11. Bergström, A. *et al.* Origins and genetic legacy of prehistoric dogs. *Science* **564**, 557–564 (2020).
12. Fages, A. *et al.* Tracking Five Millennia of Horse Management with Extensive Ancient Genome Time Series. *Cell* **177**, 1419-1435.e31 (2019).
13. Jones, M. R. *et al.* Adaptive introgression underlies polymorphic seasonal camouflage in snowshoe hares. *Science* **1358**, 1355–1358 (2018).
14. Awange, J. L. & Kyalo Kiema, J. B. Adaptive introgression enables evolutionary rescue from extreme environmental pollution. *Science* **126**, 455–457 (2019).
15. Chen, C. & Pfennig, K. S. Female toads engaging in adaptive hybridization prefer high-quality heterospecifics as mates. *Science (80-.).* **367**, 1377–1379 (2020).
16. Huang, X. *et al.* A map of rice genome variation reveals the origin of cultivated rice. *Nature* **490**, 497–501 (2012).
17. Hufford, M. B. *et al.* The Genomic Signature of Crop-Wild Introgression in Maize. *PLoS Genet.* **9**, e1003477 (2013).
18. Wang, X., Chen, L. & Ma, J. Genomic introgression through interspecific hybridization counteracts genetic bottleneck during soybean domestication. *Genome Biol.* **20**, 1–15 (2019).
19. Choi, J. Y. & Purugganan, M. D. Multiple origin but single domestication led to Oryza sativa. *G3 Genes, Genomes, Genet.* **8**, 797–803 (2018).
20. Flowers, J. M. *et al.* Cross-species hybridization and the origin of North African date palms. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 1651–1658 (2019).
21. Burgarella, C. *et al.* Adaptive introgression: An untapped evolutionary mechanism for crop adaptation. *Front. Plant Sci.* **10**, 1–17 (2019).
22. Vergauwen, D. & De Smet, I. From early farmers to Norman Borlaug — the making of modern wheat. *Curr. Biol.* **27**, R858–R862 (2017).
23. Shiferaw, B. *et al.* Crops that feed the world 10. Past successes and future challenges

to the role played by wheat in global food security. *Food Secur.* **5**, 291–317 (2013).
24. Rippke, U. *et al.* Timescales of transformational climate change adaptation in sub-Saharan African agriculture. *Nat. Clim. Chang.* **6**, 605–609 (2016).
25. Kahiluoto, H. *et al.* Decline in climate resilience of european wheat. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 123–128 (2019).
26. Liang, X. Z. *et al.* Determining climate effects on US total agricultural productivity. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E2285–E2292 (2017).
27. Asseng, S. *et al.* Rising temperatures reduce global wheat production. *Nat. Clim. Chang.* **5**, 143–147 (2015).
28. Meyer, R. S. & Purugganan, M. D. Evolution of crop species: genetics of domestication and diversification. *Nat. Rev. Genet.* **14**, 840–852 (2013).
29. Dubcovsky, J. & Dvorak, J. Genome Plasticity a Key Factor in the Success of Polyploid Wheat Under Domestication. *Science* **316**, 1862–1866 (2007).
30. He, F. *et al.* Exome sequencing highlights the role of wild-relative introgression in shaping the adaptive landscape of the wheat genome. *Nat. Genet.* **51**, 896–904 (2019).
31. Zhou, Y. *et al.* Triticum population sequencing provides insights into wheat adaptation. *Nat. Genet.* (2020). doi:10.1038/s41588-020-00722-w
32. CIMMYT. Genome of wheat, the world's most common crop. (2014).
33. FAO. Statistics Statistical Yearbook. (2020).
34. Shewry, P. R. Wheat. *Journal of Experimental Botany* **60**, 1537–1553 (2009).
35. Cooper, R. Re-discovering ancient wheat varieties as functional foods. *J. Tradit. Complement. Med.* **5**, 138–143 (2015).
36. Wang, H. *et al.* Horizontal gene transfer of Fhb7 from fungus underlies Fusarium head blight resistance in wheat. *Science* **368**, (2020).
37. Hao, M. *et al.* The Resurgence of Introgression Breeding, as Exemplified in Wheat Improvement. *Front. Plant Sci.* **11**, 1–11 (2020).
38. Huang, S. *et al.* Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the Triticum/Aegilops complex and the evolutionary history of polyploid wheat. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 8133–8138 (2002).
39. Edet, O. U., Gorafi, Y. S. A., Nasuda, S. & Tsujimoto, H. DArTseq-based analysis of genomic relationships among species of tribe Triticeae. *Sci. Rep.* **8**, 1–11 (2018).
40. Van Slageren MW. Wild wheats: a monograph of Aegilops L. and Amblyopyrum (Jaub. and Spach) Eig (Poaceae). *Agric. Univ. Wageningen, Netherlands* (1994). doi:10.2989/10220110509485863
41. Matsuoka, Y. Evolution of Polyploid Triticum Wheats under Cultivation: The Role of Domestication, Natural Hybridization and Allopolyploid Speciation in their Diversification. *Plant Cell Physiol.* **52**, 750–764 (2011).
42. Lev-Yadun, S., Gopher, A. & Abbo, S. The cradle of agriculture. *Science (New York, N.Y.)* **288**, 1602–1603 (2000).
43. Salamini, F., Özkan, H., Brandolini, A., Schäfer-Pregl, R. & Martin, W. Genetics and geography of wild cereal domestication in the near east. *Nat. Rev. Genet.* **3**, 429–441 (2002).
44. Marcussen, T. *et al.* Ancient hybridizations among the ancestral genomes of bread wheat. *Science* **345**, 1250092 (2014).
45. Salamini, F., Özkan, H., Brandolini, A., Schäfer-Pregl, R. & Martin, W. Genetics and geography of wild cereal domestication in the near east. *Nat. Rev. Genet.* **3**, 429–441 (2002).
46. Balfourier, F. *et al.* Worldwide phylogeography and history of wheat genetic diversity. *Sci. Adv.* **5**, eaav0536 (2019).
47. Harrison, R. G. & Larson, E. L. Hybridization, introgression, and the nature of species

boundaries. *J. Hered.* **105**, 795–809 (2014).
48. Keidar-Friedman, D., Bariah, I. & Kashkush, K. Genome-wide analyses of miniature inverted-repeat transposable elements reveals new insights into the evolution of the triticum-Aegilops group. *PLoS One* **13**, 1–23 (2018).
49. Bernhardt, N. *et al.* Genome-wide sequence information reveals recurrent hybridization among diploid wheat wild relatives. *Plant J.* **102**, 493–506 (2020).
50. Przewieslik-Allen, A. M. *et al.* Developing a high-throughput snp-based marker system to facilitate the introgression of traits from aegilops species into bread wheat (Triticum aestivum). *Front. Plant Sci.* **9**, 1–12 (2019).
51. Lukaszewski, A. J. *et al.* A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. *Science* **345**, (2014).
52. Van De Peer, Y., Mizrachi, E. & Marchal, K. Polyploidy The evolutionary significance of polyploidy. *Nat. Rev. Genet.* **18**, (2017).
53. Zeven AC. Polyploidy and domestication: the origin and survival of polyploids in cytotype mixtures. *Basic Life Sci* 385–407 (1980).
54. Venske, E., Dos Santos, R. S., Busanello, C., Gustafson, P. & Costa de Oliveira, A. Bread wheat: a role model for plant domestication and breeding. *Hereditas* **156**, 16 (2019).
55. Marcussen, T. *et al.* Ancient hybridizations among the ancestral genomes of bread wheat. *Science* **345**, 1250092 (2014).
56. Li, C. *et al.* Coevolution in Hybrid Genomes: Nuclear-Encoded Rubisco Small Subunits and Their Plastid-Targeting Translocons Accompanying Sequential Allopolyploidy Events in Triticum. *Mol. Biol. Evol.* 1–15 (2020). doi:10.1093/molbev/msaa158
57. El Baidouri, M. *et al.* Reconciling the evolutionary origin of bread wheat (Triticum aestivum). *New Phytol.* **213**, 1477–1486 (2017).
58. Glémin, S. *et al.* Pervasive hybridizations in the history of wheat relatives. *Sci. Adv.* **5**, (2019).
59. Li, L. F., Liu, B., Olsen, K. M. & Wendel, J. F. A re-evaluation of the homoploid hybrid origin of Aegilops tauschii, the donor of the wheat D-subgenome. *New Phytol.* **208**, 4–8 (2015).
60. Salina, E. A. *et al.* Phylogenetic reconstruction of Aegilops section Sitopsis and the evolution of tandem repeats in the diploids and derived wheat polyploids. *Genome* **49**, 1023–1035 (2006).
61. Ruban, A. S. & Badaeva, E. D. Evolution of the S-genomes in triticum-aegilops alliance: Evidences from chromosome analysis. *Front. Plant Sci.* **871**, 1–25 (2018).
62. Soltis, P. S. & Soltis, D. E. The role of hybridization in plant speciation. *Annu. Rev. Plant Biol.* **60**, 561–588 (2009).
63. Abbott, R. *et al.* Hybridization and speciation. *J. Evol. Biol.* **26**, 229–246 (2013).
64. Rodríguez, S., Maestra, B., Perera, E., Díez, M. & Naranjo, T. Pairing affinities of the B- and G-genome chromosomes of polyploid wheats with those of Aegilops speltoides. *Genome* **43**, 814–819 (2000).
65. Luo, M.-C. *et al.* The structure of wild and domesticated emmer wheat populations, gene flow between them, and the site of emmer domestication. *Theor. Appl. Genet.* **114**, 947–959 (2007).
66. Özkan, H., Willcox, G., Graner, A., Salamini, F. & Kilian, B. Geographic distribution and domestication of wild emmer wheat (Triticum dicoccoides). *Genet. Resour. Crop Evol.* **58**, 11–53 (2011).
67. Nesbitt, M. Wheat evolution: integrating archaeological and biological evidence. *Wheat Taxon. Leg. John Percival* **3**, 37–59 (2001).

68. Dvorak, J., Luo, M. C., Yang, Z. L. & Zhang, H. B. The structure of the Aegilops tauschii genepool and the evolution of hexaploid wheat. *Theor. Appl. Genet.* **97**, 657–670 (1998).
69. Singh, N. *et al.* Genomic analysis confirms population structure and identifies inter-lineage hybrids in Aegilops tauschii. *Front. Plant Sci.* **10**, 9 (2019).
70. Dvorak, J., Di Terlizzi, P., Zhang, H. B. & Resta, P. The evolution of polyploid wheats: Identification of the A genome donor species. *Genome* **36**, 21–31 (1993).
71. Wang, J. *et al.* Aegilops tauschii single nucleotide polymorphisms shed light on the origins of wheat D-genome genetic diversity and pinpoint the geographic origin of hexaploid wheat. *New Phytol.* **198**, 925–937 (2013).
72. Coyne, J. A. & Orr, A. H. Speciation. *Sunderland, MA: Sinauer.* (2004).
73. Mallet, J. Hybrid speciation. *Nature* **446**, 279–283 (2007).
74. Cheng, H. *et al.* Frequent intra- and inter-species introgression shapes the landscape of genetic variation in bread wheat. *Genome Biol.* **20**, (2019).
75. Dvorak, J. and Luo, M. . *THE Wheat Taxonomy : the legacy of John Percival*. (2001).
76. Kuckuck, H. On the origin of Triticum carthlicum (= T. persicum). *Wheat Inf. Serv.* **50**, 1–5 (1979).
77. Mac Key, J. Species relationship in Triticum. *Hereditas* **Suppl. 2**, 237–275 (1966).
78. Blatter, R. H. E., Jacomet, S. & Schlumbaum, A. Spelt-specific alleles in HMW glutenin genes from modern and historical European spelt (Triticum spelta L.). *Theor. Appl. Genet.* **104**, 329–337 (2002).
79. Yan, Y. *et al.* HMW and LMW glutenin alleles among putative tetraploid and hexaploid European spelt wheat (Triticum spelta L.) progenitors. *Theor. Appl. Genet.* **107**, 1321–1330 (2003).
80. Blatter, R. H. E., Jacomet, S. & Schlumbaum, A. About the origin of European spelt (Triticum spelta L.): allelic differentiation of the HMW Glutenin B1-1 and A1-2 subunit genes. *Theor. Appl. Genet.* **108**, 360–367 (2004).
81. Chen, Q. *et al.* Evolutionary history of Triticum petropavlovskyi Udacz. et Migusch. inferred from the sequences of the 3-phosphoglycerate kinase gene. *PLoS One* **8**, 19–21 (2013).
82. Nyine, M. *et al.* Genomic patterns of introgression in interspecific populations created by crossing wheat with its wild relative. *G3 Genes, Genomes, Genet.* **10**, 3651–3661 (2020).
83. Scott, M. F. *et al.* A 3,000-year-old Egyptian emmer wheat genome reveals dispersal and domestication history. *Nat. Plants* **5**, 1120–1128 (2019).
84. Janzen, G. M., Wang, L. & Hufford, M. B. The extent of adaptive wild introgression in crops. *New Phytol.* **221**, 1279–1288 (2018).
85. Hedrick, P. W. Adaptive introgression in animals: Examples and comparison to new mutation and standing variation as sources of adaptive variation. *Mol. Ecol.* **22**, 4606–4618 (2013).
86. Rochus, C. M. *et al.* Revealing the selection history of adaptive loci using genome-wide scans for selection: An example from domestic sheep. *BMC Genomics* **19**, 1–17 (2018).
87. Racimo, F., Sankararaman, S., Nielsen, R. & Huerta-Sánchez, E. Evidence for archaic adaptive introgression in humans. *Nat. Rev. Genet.* **16**, 359–371 (2015).
88. Flowers, J. M. *et al.* Cross-species hybridization and the origin of North African date palms. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 1651–1658 (2019).
89. Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G. & François, O. Fast and efficient estimation of individual ancestry coefficients. *Genetics* **196**, 973–983 (2014).
90. Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. Bayesian phylogenetics

with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–1973 (2012).
91. Than, C., Ruths, D. & Nakhleh, L. PhyloNet: A software package for analyzing and reconstructing reticulate evolutionary relationships. *BMC Bioinformatics* **9**, 1–16 (2008).
92. Cornuet, J. M. *et al.* DIYABC v2.0: A software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* **30**, 1187–1189 (2014).
93. Roda, F., Mendes, F. K., Hahn, M. W. & Hopkins, R. Genomic evidence of gene flow during reinforcement in Texas Phlox. *Mol. Ecol.* **26**, 2317–2330 (2017).
94. Patterson, N. *et al.* Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
95. Martin, S. H., Davey, J. W. & Jiggins, C. D. Evaluating the use of ABBA-BABA statistics to locate introgressed loci. *Mol. Biol. Evol.* **32**, 244–257 (2015).
96. Blischak, P. D., Chifman, J., Wolfe, A. D. & Kubatko, L. S. HyDe: A python package for genome-scale hybridization detection. *Syst. Biol.* **67**, 821–829 (2018).
97. Pease, J. B. & Hahn, M. W. Detection and Polarization of Introgression in a Five-Taxon Phylogeny. *Syst. Biol.* **64**, 651–662 (2015).
98. Peter, B. M. Admixture, Population Structure and F-statistics. *Genetics* (2016). doi:10.1534/genetics.XXX.XXXXXX
99. Vernot, B. *et al.* Excavating Neandertal and Denisovan DNA from the genomes of Melanesian individuals. *Science* **352**, 235–239 (2016).
100. Racimo, F., Marnetto, D. & Huerta-Sánchez, E. Signatures of archaic adaptive introgression in present-day human populations. *Mol. Biol. Evol.* **34**, 296–317 (2017).
101. Sattler, M. C., Carvalho, C. R. & Clarindo, W. R. The polyploidy and its key role in plant breeding. *Planta* **243**, 281–296 (2016).
102. Otto, S. P. The Evolutionary Consequences of Polyploidy. *Cell* **131**, 452–462 (2007).
103. Comai, L. The advantages and disadvantages of being polyploid. *Nat. Rev. Genet.* **6**, 836–846 (2005).
104. Luo, M.-C. *et al.* Genome sequence of the progenitor of the wheat D genome Aegilops tauschii. *Nature* **551**, 498 (2017).
105. Wang, L. *et al.* Molecular Parallelism Underlies Convergent Highland Adaptation of Maize Landraces. *bioRxiv* 2020.07.31.227629 (2020).
106. Grant, P. R. G. and B. R. Phenotypic and Genetic Effects of Hybridization in Darwin's Finches. 297–316 (1994).
107. Ellstrand, N. C. Is gene flow the most important evolutionary force in plants? *Am. J. Bot.* **101**, 737–753 (2014).
108. Meyer, R. S. *et al.* Domestication history and geographical adaptation inferred from a SNP map of African rice. *Nat. Genet.* **48**, 1083–1088 (2016).
109. Rabinovich, S. V. Importance of wheat-rye translocations for breeding modern cultivars of Triticum aestivum L. *Euphytica* **100**, 323–340 (1998).
110. Danilova, T. V., Poland, J. & Friebe, B. Production of a complete set of wheat–barley group-7 chromosome recombinants with increased grain β-glucan content. *Theor. Appl. Genet.* **132**, 3129–3141 (2019).
111. Crow, B. Y. J. F. SOME POSSIBILITIES FOR MEASURING SELECTION INTENSITIES IN MAN Author (s): JAMES F . CROW. *Hum. Biol.* **61**, 763–775 (1989).
112. Kim, B. Y., Huber, C. D. & Lohmueller, K. E. Deleterious variation shapes the genomic landscape of introgression. *PLoS Genet.* **14**, 1–30 (2018).
113. Wang, L. *et al.* The interplay of demography and selection during maize domestication

and expansion. *Genome Biol.* **18**, 1–13 (2017).

114. Hardigan, M. A. *et al.* Genome diversity of tuber-bearing Solanum uncovers complex evolutionary history and targets of domestication in the cultivated potato. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E9999–E10008 (2017).
115. Lovell, J. T. *et al.* Genomic mechanisms of climate adaptation in polyploid bioenergy switchgrass. *Nature* **590**, (2021).
116. Todesco, M. *et al.* Massive haplotypes underlie ecotypic differentiation in sunflowers. *Nature* (2020). doi:10.1101/790279
117. Birchler, J. A. & Veitia, R. A. Gene balance hypothesis: Connecting issues of dosage sensitivity across biological disciplines. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 14746–14753 (2012).
118. Schumer, M. *et al.* Natural selection interacts with recombination to shape the evolution of hybrid genomes. *Science* **660**, 656–660 (2018).
119. Johnson, N. A. Speciation: Dobzhansky-Muller incompatibilities, dominance and gene interactions. *Trends Ecol. Evol.* **15**, 480–482 (2000).
120. Wu, C. I. & Ting, C. T. Genes and speciation. *Nat. Rev. Genet.* **5**, 114–122 (2004).
121. Chen, Z. J. & Ni, Z. Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. *Bioessays* **23**, 1–7 (2011).
122. Cheng, F. *et al.* Gene retention, fractionation and subgenome differences in polyploid plants. *Nat. Plants* **4**, 258–268 (2018).
123. Innan, H. & Kondrashov, F. The evolution of gene duplications: classifying and distinguishing between models. *Nat. Rev. Genet.* **11**, (2010).
124. Timothy Rabanus-Wallace, M. *et al.* Chromosome-scale genome assembly provides insights into rye biology, evolution, and agronomic potential. *Nature Genetics* (2019). doi:10.1101/2019.12.11.869693
125. Li, A., Liu, D., Yang, W., Kishii, M. & Mao, L. Synthetic Hexaploid Wheat: Yesterday, Today, and Tomorrow. *Engineering* **4**, 552–558 (2018).
126. Conant, G. C., Birchler, J. A. & Pires, J. C. Dosage, duplication, and diploidization: Clarifying the interplay of multiple models for duplicate gene evolution over time. *Curr. Opin. Plant Biol.* **19**, 91–98 (2014).
127. Zhang, Z. *et al.* Duplication and partitioning in evolution and function of homoeologous Q loci governing domestication characters in polyploid wheat. *PNAS* **109**, 1353 (2012).
128. Singh, N. K., Shepherd, K. W. & McIntosh, R. A. Linkage mapping of genes for resistance to leaf, stem and stripe rusts and ω-secalins on the short arm of rye chromosome 1R. *Theor. Appl. Genet.* **80**, 609–616 (1990).
129. Burnett, C. J., Lorenz, K. J. & Carver, B. F. Effects of the 1B/1R translocation in wheat on composition and properties of grain and flour. *Euphytica* **86**, 159–166 (1995).
130. Simons, K. J. *et al.* Molecular characterization of the major wheat domestication gene Q. *Genetics* **172**, 547–555 (2006).
131. Sood, S., Kuraparthy, V., Bai, G. & Gill, B. S. The major threshability genes soft glume (sog) and tenacious glume (Tg), of diploid and polyploid wheat, trace their origin to independent mutations at non-orthologous loci. *Theor. Appl. Genet.* **119**, 341–351 (2009).
132. Martin, S. H., Davey, J. W., Salazar, C. & Jiggins, C. D. Recombination rate variation shapes barriers to introgression across butterfly genomes. *PLoS Biol.* **17**, 1–28 (2019).
133. Dreissig, S. *et al.* Natural variation in meiotic recombination rate shapes introgression patterns in intraspecific hybrids between wild and domesticated barley. *New Phytol.* **228**, 1852–1863 (2020).

134. Rieseberg, L. H., Whitton, J. & Gardner, K. Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. *Genetics* **152**, 713–727 (1999).
135. Wolf, D. E. *et al.* Predicting the Risk of Extinction through Hybridization. *Soc. Conserv. Biol.* **15**, 1039–1053 (2001).
136. Merotto, A. *et al.* Evolutionary and social consequences of introgression of nontransgenic herbicide resistance from rice to weedy rice in Brazil. *Evol. Appl.* **9**, 837–846 (2016).
137. Todesco, M. *et al.* Hybridization and extinction. *Evol. Appl.* **9**, 892–908 (2016).
138. Harris, K. & Nielsen, R. The genetic cost of neanderthal introgression. *Genetics* **203**, 881–891 (2016).
139. Rotival, M. & Quintana-Murci, L. Functional consequences of archaic introgression and their impact on fitness. *Genome Biol.* **21**, 19–22 (2020).
140. Harris, K. & Nielsen, R. Inferring Demographic History from a Spectrum of Shared Haplotype Lengths. *PLoS Genet.* **9**, (2013).
141. Seixas, F. A., Boursot, P. & Melo-Ferreira, J. The genomic impact of historical hybridization with massive mitochondrial DNA introgression. *Genome Biol.* **19**, 1–20 (2018).
142. Petr, M., Pääbo, S., Kelso, J. & Vernot, B. Limits of long-term selection against Neandertal introgression. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 1639–1644 (2019).
143. Zhang, W. *et al.* Molecular characterization of durum and common wheat recombinant lines carrying leaf rust resistance (Lr19) and yellow pigment (Y) genes from Lophopyrum ponticum. *Theor. Appl. Genet.* **111**, 573–582 (2005).
144. Li, Z.-S., Li, B., Zheng, Q. & Li, H. Review and New Progress in Wheat Wide Hybridization for Improving the Resistance to Biotic and Abiotic Stresses. in *Advances in Wheat Genetics: From Genome to Field* 377–385 (Springer Japan, 2015). doi:10.1007/978-4-431-55675-6_43
145. Wilson, A. S. On wheat and rye hybrids. *Trans. Proc. Bot. Soc. Edinburgh* 286–288 (1876). doi:10.1007/978-3-662-10933-5_7
146. Brown, J. & Caligari, P. D. S. *An Introduction to Plant Breeding*. **43**, (2008).
147. Sears, E. R. “Chromosome engineering in wheat,” in Proceedings ofthe 4th Stadler Genetics Symposium. **13**, (1972).
148. Sanei, M., Pickering, R., Kumke, K., Nasuda, S. & Houben, A. Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. *Proc. Natl. Acad. Sci. U. S. A.* **108**, (2011).
149. Hao, M. *et al.* QTug.sau-3B Is a major quantitative trait locus for wheat hexaploidization. *G3 Genes, Genomes, Genet.* **4**, 1943–1953 (2014).
150. Harlan, J. R. & Wet, J. M. J. de. *Toward a Rational Classification of Cultivated*. **20**, (1971).
151. Halloran, G. M., Ogbonnaya, F. C. & Lagudah, E. S. Triticum (Aegilops) tauschii in the natural and artificial synthesis of hexaploid wheat. *Aust. J. ofAgricultural Res.* (2008). doi:10.1071/AR07352
152. Dubcovsky, J., Lukaszewski, A. J., Echaide, M., Antonelli, E. F. & Porter, D. R. Molecular characterization of two Triticum speltoides interstitial translocations carrying leaf rust and greenbug resistance genes. *Crop Sci.* **38**, 1655–1660 (1998).
153. Przewieslik-allen, A. M. *et al.* The role of gene flow and chromosomal instability in shaping the bread wheat genome. *Nat. Plants* (2021). doi:10.1038/s41477-020-00845-2
154. Fluch, S. *et al.* Sequence composition and gene content of the short arm of rye (Secale cereale) chromosome 1. *PLoS One* **7**, 4–9 (2012).

155. Qi, L. L. *et al.* A novel Robertsonian translocation event leads to transfer of a stem rust resistance gene (Sr52) effective against race Ug99 from Dasypyrum villosum into bread wheat. *Theor. Appl. Genet.* **123**, 159–167 (2011).
156. Rahmatov, M. *et al.* A new 2DS·2RL Robertsonian translocation transfers stem rust resistance gene Sr59 into wheat. *Theor. Appl. Genet.* **129**, 1383–1392 (2016).
157. Hao, M. *et al.* Introgression of powdery mildew resistance gene pm56 on rye chromosome arm 6rs into wheat. *Front. Plant Sci.* **9**, 1–8 (2018).
158. Liu, W. *et al.* Development and characterization of a compensating wheat-Thinopyrum intermedium Robertsonian translocation with Sr44 resistance to stem rust (Ug99). *Theor. Appl. Genet.* **126**, 1167–1177 (2013).
159. Danilova, T. V., Friebe, B., Gill, B. S., Poland, J. & Jackson, E. Development of a complete set of wheat–barley group-7 Robertsonian translocation chromosomes conferring an increased content of β-glucan. *Theor. Appl. Genet.* **131**, 377–388 (2018).
160. Hao, C. *et al.* Resequencing of 145 Landmark Cultivars Reveals Asymmetric Sub-genome Selection and Strong Founder Genotype Effects on Wheat Breeding in China. *Mol. Plant* **13**, 1733–1751 (2020).
161. Wang, H. *et al.* Sympatric speciation of wild emmer wheat driven by ecology and chromosomal rearrangements. *Proc. Natl. Acad. Sci.* 201920415 (2020). doi:10.1073/pnas.1920415117
162. Guo, W. *et al.* Origin and adaptation to high altitude of Tibetan semi-wild wheat. *Nat. Commun.* **11**, 1–12 (2020).
163. Jordan, K. W. *et al.* A haplotype map of allohexaploid wheat reveals distinct patterns of selection on homoeologous genomes. *Genome Biol.* **16**, 1–18 (2015).
164. Wang, S. *et al.* Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnol. J.* **12**, 787–796 (2014).
165. Chen, K., Wang, Y., Zhang, R., Zhang, H. & Gao, C. CRISPR/Cas Genome Editing and Precision Plant Breeding in Agriculture. *Annu. Rev. Plant Biol.* **70**, 667–697 (2019).
166. Zhu, H., Li, C. & Gao, C. Applications of CRISPR–Cas in agriculture and plant biotechnology. *Nat. Rev. Mol. Cell Biol.* **21**, 661–677 (2020).
167. Ramírez-González, R. H. *et al.* The transcriptional landscape of polyploid wheat. *Science* **361**, eaar6089 (2018).
168. Gardiner, L.-J. *et al.* A genome-wide survey of DNA methylation in hexaploid wheat. *Genome Biol.* **16**, 273 (2015).
169. Li, Z. *et al.* The bread wheat epigenomic map reveals distinct chromatin architectural and evolutionary features of functional genetic elements. *Genome Biol.* **20**, 1–16 (2019).
170. Jordan, K. W., He, F., De Soto, M. F., Akhunova, A. & Akhunov, E. Differential chromatin accessibility landscape reveals structural and functional features of the allopolyploid wheat chromosomes. *Genome Biol.* **21**, 1–30 (2020).
171. Wang, M. *et al.* An atlas of wheat epigenetic regulatory elements reveals subgenome-divergence in the regulation of development and stress responses An atlas of wheat epigenetic regulatory elements reveals subgenome-divergent regulation of development and stress responses. *Plant Cell* (2021).
172. Cullot, G. *et al.* CRISPR-Cas9 genome editing induces megabase-scale chromosomal truncations. *Nat. Commun.* **10**, 1–14 (2019).