

Evolutionary History of *Triticum petropavlovskyi* Udacz. et Migusch. Inferred from the Sequences of the 3-Phosphoglycerate Kinase Gene

Qian Chen^{1,2,3}, Hou-Yang Kang^{1,3}, Xing Fan¹, Yi Wang¹, Li-Na Sha¹, Hai-Qin Zhang¹, Mei-Yu Zhong¹, Li-Li Xu¹, Jian Zeng³, Rui-Wu Yang⁴, Li Zhang⁴, Chun-Bang Ding⁴, Yong-Hong Zhou^{1,2*}

1 Triticeae Research Institute, Sichuan Agricultural University, Sichuan, People's Republic of China, **2** Key Laboratory of Crop Genetic Resources and Improvement, Ministry of Education, Sichuan Agricultural University, Sichuan, People's Republic of China, **3** College of Resources and Environment, Sichuan Agricultural University, Sichuan, People's Republic of China, **4** College of Biology and Science, Sichuan Agricultural University, Sichuan, People's Republic of China

Abstract

Single- and low-copy genes are less likely to be subject to concerted evolution. Thus, they are appropriate tools to study the origin and evolution of polyploidy plant taxa. The plastid 3-phosphoglycerate kinase gene (*Pgk-1*) sequences from 44 accessions of *Triticum* and *Aegilops*, representing diploid, tetraploid, and hexaploid wheats, were used to estimate the origin of *Triticum petropavlovskyi*. Our phylogenetic analysis was carried out on exon-intron, exon and intron sequences, using maximum likelihood, Bayesian inference and haplotype networking. We found the D genome sequences of *Pgk-1* genes from *T. petropavlovskyi* are similar to the D genome orthologs in *T. aestivum*, while their relationship with *Ae. tauschii* is more distant. The A genome sequences of *T. petropavlovskyi* group with those of *T. polonicum*, but its *Pgk-1* B genome sequences to some extent diverge from those of other species of *Triticum*. Our data do not support for the origin of *T. petropavlovskyi* either as an independent allopolyploidization event between *Ae. tauschii* and *T. polonicum*, or as a monomendelian mutation in *T. aestivum*. We suggest that *T. petropavlovskyi* originated via spontaneous introgression from *T. polonicum* into *T. aestivum*. The dating of this introgression indicates an age of 0.78 million years; a further mutation event concerning the B genome occurred 0.69 million years ago.

Citation: Chen Q, Kang H-Y, Fan X, Wang Y, Sha L-N, et al. (2013) Evolutionary History of *Triticum petropavlovskyi* Udacz. et Migusch. Inferred from the Sequences of the 3-Phosphoglycerate Kinase Gene. PLoS ONE 8(8): e71139. doi:10.1371/journal.pone.0071139

Editor: Tianzhen Zhang, Nanjing Agricultural University, China

Received: May 17, 2013; **Accepted:** July 1, 2013; **Published:** August 19, 2013

Copyright: © 2013 Chen et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was financially supported by the National Natural Science Foundation of China (31101151), the National Transgenic Major Project (2011zx08009001), the National High-tech R&D Program of China (863 Program) (2011AA100103-02), the Special Fund for Agro-Scientific Research in the Public Interest of China (201003021), and the Science and Technology Bureau and Education Bureau of Sichuan Province, China. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: zhouyh@sicau.edu.cn

These co-first authors contributed equally to this work.

Introduction

In the Xinjiang region of China, *Triticum* species are abundant. The Xinjiang rice wheat (*Triticum petropavlovskyi* Udacz. et Migusch.), known as 'Daosuimai' or rice-head wheat, is one of the Chinese endemic wheat landraces, together with the Sichuan white wheat complex (*T. aestivum* L.), Tibetan weedrace (*T. aestivum* ssp. *tibetanum* Shao) and Yunnan hulled wheat (*T. aestivum* ssp. *yunnanense* King) [1]. Based on chromosome pairing, morphology, eco-geographical origin and RFLP analysis, the Xinjiang rice wheat is distinct from other Chinese endemic wheat landraces [2–5].

The origin of *T. petropavlovskyi* has been discussed for decades. Gorsky [6] analyzed the morphology of Xinjiang rice wheat, and suggested that it was a mutated form of the tetraploid *Triticum polonicum* L.. However, Udachin and Miguschova [7] discovered that the Xinjiang rice wheat is not tetraploid but hexaploid, and named it *T. petropavlovskyi* Udacz. et Migusch. The chromosomal constitutions of the Xinjiang rice wheat is AABBDD [2,8–10]. Dorofeev et al. [11] hypothesized that *T. petropavlovskyi* could be the result of spontaneous hybridization between *T. aestivum* and *T.*

polonicum. Several previous studies indicated that the genes supporting a long glume in *T. polonicum* and *T. petropavlovskyi* were allelic and located on the long arm of chromosome 7A [12–14], agreeing with the hypothesis of spontaneous hybridization. Yen et al. [15] studied the natural distribution in Xinjiang of *Aegilops tauschii*, and Yang et al. [16] and Liu et al. [17] reported a similar distribution for a dwarfing accession of *T. polonicum* and suggested that *T. petropavlovskyi* is derived from a hybridization event between *T. polonicum* and *Ae. tauschii*. However, Efremova et al. [18] maintained that *T. petropavlovskyi* originated from *T. aestivum* through spontaneous mutation.

Despite prior intensive research, the origin of *T. petropavlovskyi* is still uncertain, and three hypotheses have been proposed: (1) *T. petropavlovskyi* is an independent species is derived from a natural hybridization event between *T. polonicum* and *Ae. tauschii* [10,15,19–21]; (2) *T. petropavlovskyi* is a natural cross or backcross between *T. polonicum* and *T. aestivum* [2,11,12,22,23]; and (3) *T. petropavlovskyi* is a monogenic mutant of *T. aestivum* [18,24]. In a recently study, Kang et al. [25] created the synthetic hexaploid wheat (SHW-DPW) between *T. polonicum* from Xinjiang and *Ae. tauschii*: its spike

Table 1. Plants used in this study.

Species	Genome	Accession	Origin	Abbrev.	GenBank Ac. No.
<i>Triticum urartu</i> Thum. ex Gandil.	A ^u	TA763	Lebanon	TUR63A	AF343474
<i>Aegilops bicornis</i> (Forskal) Jaub. et Spach.	S ^b	TA1954	Egypt	AEB954S	AF343485
<i>Aegilops longissima</i> Schweinf. et Muschl.	S ^l	TA1912	Israel	AEL912S	AF343487
<i>Aegilops searsii</i> Feldman et Kislev	S ^s	TA2355	Israel	AES355S	AF343489
<i>Aegilops sharonensis</i> Eig	S ^{sh}	TA2065	Turkey	AES065S	AF343486
<i>Aegilops speltoides</i> Tausch	S	TA2368	Turkey	AES368S	AF343483
<i>Aegilops speltoides</i> var. <i>ligustica</i> (Savign.) Fiori	S	TA1770	Iraq	AEL770S	AF343484
<i>Aegilops tauschii</i> Cosson	D	AS60	Middle East	AET60D	JQ327050
		TA1691	Unkown	AET691D	AF343479
<i>Triticum polonicum</i> L.	AB	AS302	Xinjiang, China	TPO302A	JQ327101
				TPO302B	JQ327102
		AS304	Xinjiang, China	TPO304A	JQ327088
				TPO304B	JQ327089
		PI42209	Australia	TPO209A	JQ327096
				TPO209B	JQ327097
<i>Triticum turgidum</i> L.	AB	AS2233	Xinjiang, China	TUR233A	JQ327113
				TUR233B	JQ327114
		AS2277	Xinjiang, China	TUR277A	JQ327077
				TUR277B	JQ327078
<i>Triticum durum</i> Desf.	AB	AS2349	Xinjiang, China	TDU349A	JQ327115
				TDU349B	JQ327116
<i>Triticum durum</i> Desf. cv. Langdon	AB	LDN	USA	TDULA	JQ327057
				TDULB	JQ327058
<i>Triticum turanicum</i> Jakubz.	AB	AS2229	Xinjiang, China	TTU229A	JQ327109
				TTU229B	JQ327110
		AS2279	Xinjiang, China	TTU279A	JQ327111
				TTU279B	JQ327112
<i>Triticum dicoccoides</i> (Koern. ex Aschers. et Graeb.) Schweinf.	AB	TA51	Israel	TDI51A	AF343481
				TDI51B	AF343476
		AS838	Xinjiang, China	TDI838A	JQ327075
				TDI838B	JQ327076
<i>Triticum carthlicum</i> Nevski (syn. <i>T. persicum</i> Vav.)AB		PI532494	Kars, Turkey	TCA494A	JQ327065
				TCA494B	JQ327066
		PI532509	Xinjiang, China	TCA509A	JQ327073
				TCA509B	JQ327074
<i>Triticum timopheevii</i> (Zhuk.) Zhuk.	AG	TA2	Armenia	TTI2A	AF343477
				TTI2G	AF343488
		PI94761	Georgia, USA	TTI761A	JQ327126
				TTI761G	JQ327127
<i>Triticum petropavlovskyi</i> Udacz. et Migusch.	ABD	AS358	Xinjiang, China	TPE358A	JQ327090
				TPE358B	JQ327091
				TPE358D	JQ327092
		AS359	Xinjiang, China	TPE359A	JQ327103
				TPE359B	JQ327104
				TPE359D	JQ327105
		AS360	Xinjiang, China	TPE360A	JQ327106
				TPE360B	JQ327107
				TPE360D	JQ327108
<i>Triticum aestivum</i> L. ssp. <i>tibetanum</i> Shao	ABD	AS1026	Xizang, China	TTB1026A	JQ327123

Table 1. Cont.

Species	Genome	Accession	Origin	Abbrev.	GenBank Ac. No.
				TTB1026B	JQ327124
				TTB1026D	JQ327125
		AS1027	Xizang, China	TTB1027A	JQ327062
				TTB1027B	JQ327063
				TTB1027D	JQ327064
<i>Triticum aestivum</i> L. ssp. <i>yunnanense</i> King	ABD	AS331	Yunnan, China	TYU331A	JQ327131
				TYU331B	JQ327132
				TYU331D	JQ327133
		AS338	Yunnan, China	TYU338A	JQ327085
				TYU338B	JQ327086
				TYU338D	JQ327087
		AS343	Yunnan, China	TYU343A	JQ327128
				TYU343B	JQ327129
				TYU343D	JQ327130
<i>Triticum sphaerococcum</i> Perciv.	ABD	PI70711	Iraq	TSP711A	JQ327117
				TSP711B	JQ327118
				TSP711D	JQ327119
		PI115818	Punjab, India	TSP818A	JQ327093
				TSP818B	JQ327094
				TSP818D	JQ327095
<i>Triticum macha</i> Dekapr. et Menabde.	ABD	PI278660	UK	TMA660A	JQ327082
				TMA660B	JQ327083
				TMA660D	JQ327084
<i>Triticum spelta</i> L.	ABD	PI347852	Switzerland	TPL852A	JQ327098
				TPL852B	JQ327099
				TPL852D	JQ327100
		PI347858	Switzerland	TPL858A	JQ327120
				TPL858B	JQ327121
				TPL858D	JQ327122
<i>Triticum compactum</i> Host	ABD	PI124298	Unknown	TCO298A	JQ327070
				TCO298B	JQ327071
				TCO298D	JQ327072
		PI352299	Switzerland	TCO299A	JQ327067
				TCO299B	JQ327068
				TCO299D	JQ327069
<i>Triticum aestivum</i> L. cv. Chinese Spring	ABD	CS	Sichuan, China	TCHSA	JQ327051
				TCHSB	JQ327052
				TCHSD	JQ327053
<i>Triticum aestivum</i> L. cv. Chuannong-16	ABD	CN16	Sichuan, China	TCN16A	JQ327054
				TCN16B	JQ327055
				TCN16D	JQ327056
<i>Triticum aestivum</i> L. cv. J-11	ABD	J-11	Sichuan, China	TJ11A	JQ327079
				TJ11B	JQ327080
				TJ11D	JQ327081
Synthetic hexaploid wheat	ABD	SHW-DPW		SHWDA	JQ327059
				SHWDB	JQ327060
				SHWDD	JQ327061
<i>Psathyrostachys juncea</i> (Fischer) Nevski	Ns	PI222050	Afghanistan	PJU050N	FJ711031

The Genebank with AF numbers are from Huang et al. [25], those with JQ numbers have been assigned in this study.
doi:10.1371/journal.pone.0071139.t001

Table 2. Parameters derived from *Pgk-1* sequence alignments.

	Total sites	Variable characters	Conversed characters	Informative characters
Exon+Intron	1466	174	1292	91
Exon	894	88	806	36
Intron	572	107	378	80

doi:10.1371/journal.pone.0071139.t002

morphology was similar to *T. petropavlovskyi*. However, a comparison of SHW-DPW to *T. petropavlovskyi*, *T. polonicum* and related species by the phylogenetic analysis of the *Acc-1* gene indicated that *T. petropavlovskyi* originated from the cross between *T. polonicum* from Xinjiang and exotic landraces of *T. aestivum* [26]. This finding contradicts the morphology-based conclusion of preceding study. Previous works based on different methods, including cytology, morphology and nuclear markers, have failed in identifying the origin of *T. petropavlovskyi*. Furthermore, no molecular-clock has been reported to examine the timing of its origin.

Single- and low-copy nuclear genes, being less susceptible to concerted evolution [27–29], are useful in phylogenetic study [30–33] as well as in the identification of parents of allopolyploidy taxa [26,34–36]. Genes such as acetyl-CoA carboxylase 1 (*Acc-1*) [30], disrupted meiotic cDNA 1 (*DMCI*) and translation elongation factor G (*EF-G*) [37] have been particularly useful in elucidating the phylogenesis of *Triticum-Aegilops* species. The plastid 3-phosphoglycerate kinase (PGK) gene, *Pgk-1*, is a single copy nuclear gene in diploid species of the Triticeae; it is frequently considered to be superior to the *Acc-1* gene for assessing the evolutionary history of polyploid wheats, because the *Pgk-1* gene has more parsimony informative sites than the *Acc-1* gene [31,38]. For allotetraploid and allohexaploid species with two or three copies of genes present as single copies in diploid ancestors, the *Pgk-1* gene can both elucidate the phylogenetic relationships of such polyploid as well as potential progenitors [30,31,39,40].

In this study, we sequenced and analyzed the single-copy nuclear *Pgk-1* gene in the following taxa: *T. petropavlovskyi*, SHW-DPW (synthetic hexaploid wheat between *T. polonicum* and *Ae. tauschii*), and the hypothetical *Triticum* and *Aegilops* progenitors of *T. petropavlovskyi* to reveal their phylogenetic relationships and to explore both the origin of *T. petropavlovskyi* and its divergence time from other taxa.

Materials and Methods

Plant materials

The species, genomic constitutions, origin, GenBank accessions, and sources of the taxa are listed in Table 1. The sequences of *pgk-1* gene of the accessions with TA numbers were obtained from the GenBank database; the rest of the species considered for the sequences are reported here for the first time.

The accessions with PI and AS numbers were kindly provided by the American National Plant Germplasm System (Pullman, Washington, USA) and the Triticeae Research Institute, Sichuan Agricultural University, China, respectively. The artificial synthetic amphiploid of *Triticum polonicum* and *Aegilops tauschii* (SHW-DPW) was produced by Kang et al. [25]. The plants and voucher specimens have been deposited at Herbarium of Triticeae

Research Institute, Sichuan Agricultural University, China (SAUTI).

DNA extraction, amplification and sequencing

DNA was extracted from fresh leaves of single plants, following a standard CTAB (cetyltrimethylammonium bromide) protocol [41]. For amplification of the *Pgk-1* gene, a pair of *Pgk-1*-specific primers, PPF1 (5'-CACCTGGGTCGTCCTAAGGGTGT-3') and PPR1 (5'-ACCACCAGTTGTGTTGTGGCTCAT-3'), was used [31]. Polymerase chain reactions (PCR) were performed in a GeneAmp 9700 Thermal Cycler (Applied Biosystems Inc., California, USA) according to the following cycling program: initial denaturation at 94°C for 5 min; 35 cycles of 94°C for 30 s, 56°C for 30 s, 68°C for 5 min; followed by a final elongation period at 68°C for 10 min. A final volume of 50 µl for each PCR reaction was prepared, containing 0.5 µg of genomic DNA, 10× reaction buffer, 1.5 mM of each primer, 2.5 mM of each dNTP, 2.5 mM MgCl₂, 2 units of high-fidelity ExTaq DNA polymerase (Takara Biotechnology Co. Ltd., Dalian, China).

1.0% agarose gel was used to estimate the size of the amplification products, which were purified using the EZNATN gel extraction kit (Omega Bio-Tech, Georgia, USA) and stored in 30 µl TE buffer. The purified products were cloned into the pMD19-T vector (Takara) according to the manufacturer's instructions. Cloning of PCR amplifications from single-copy nuclear genes from allopolyploid species should isolate homoeologous sequences from each nuclear genome [42,43]. For the hexaploid *Triticum* species, the A, B and D genomes homoeologous sequences of *Pgk-1* gene were isolated, and the A and B genomes homoeologous sequences were separated for the tetraploid *Triticum*. The cloned PCR products were commercially sequenced on both strands by the Beijing Genomics Institute (BGI, Shenzhen, China). All the sequences used in the phylogenetic analysis were derived from at least five independent clones.

Alignments and phylogenetic analysis

Multiple sequences were aligned using Clustal X with default parameters, followed by manual adjustment to minimize gaps [44]. In an initial phylogenetic analysis, if all sequences used for alignment derived from independent clones, formed a monophyletic group, then only one sequence was used later on. Distinct sequences derived from single accessions mapping different clades were all included in the phylogenetic analysis. Nucleotide frequencies, transition/transversion ration, and variability in different regions of the sequences were examined by MEGA 5.0 [45].

Three data matrixes, including exon+intron data (the target *Pgk-1* gene sequences), exon data and intron data, were used separately to implement phylogenetic analyses. Phylogenetic trees were

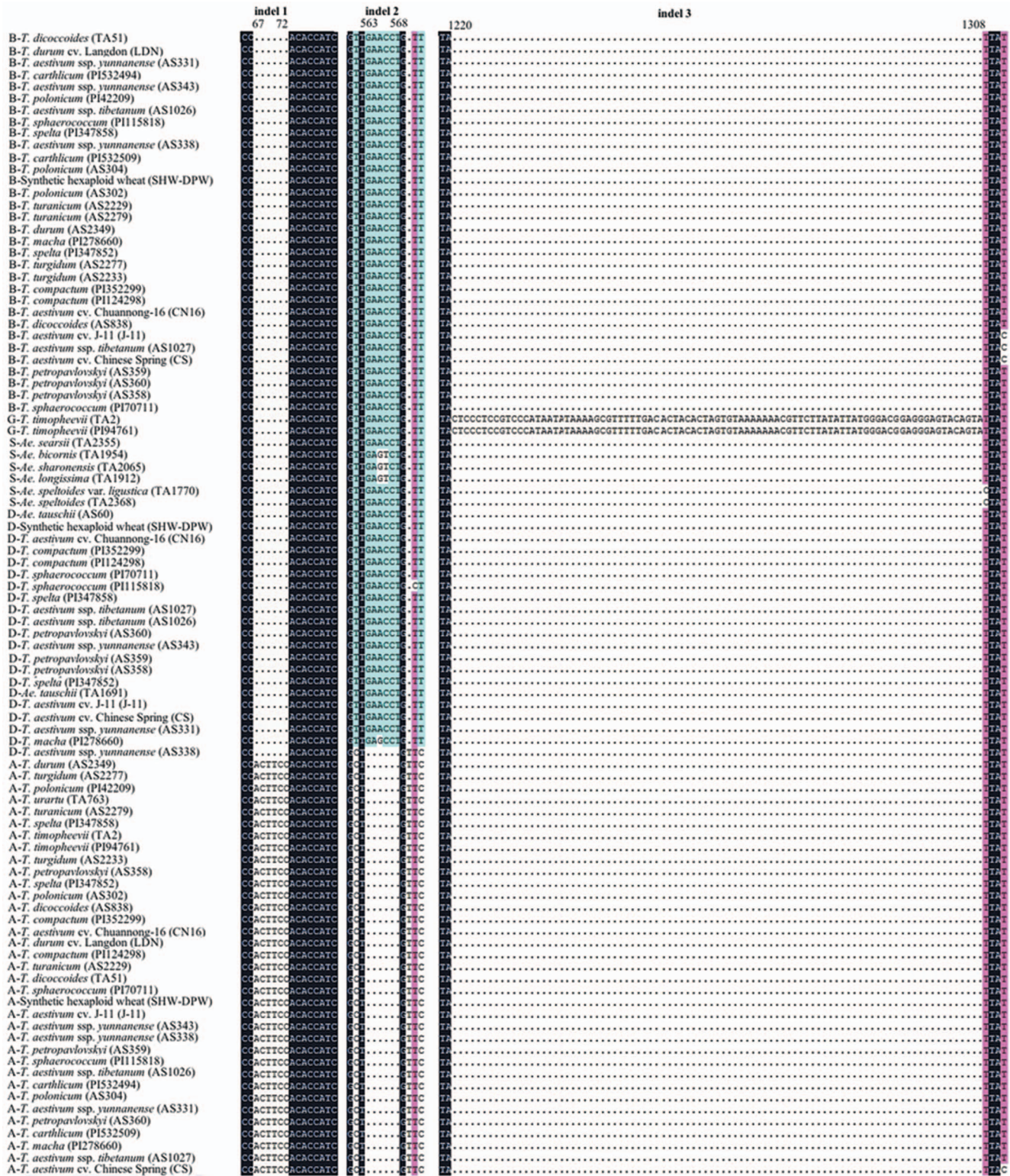


Figure 1. Maximum-likelihood tree from the exon+intron sequences of the *Pgl-1* gene of *T. petropavlovskiyi* and its related species. Numbers above nodes are bootstrap values >50% numbers below nodes are posterior probability values >90%. Genome composition, species name and accession number/cultivar name are indicated for each taxon. doi:10.1371/journal.pone.0071139.g001

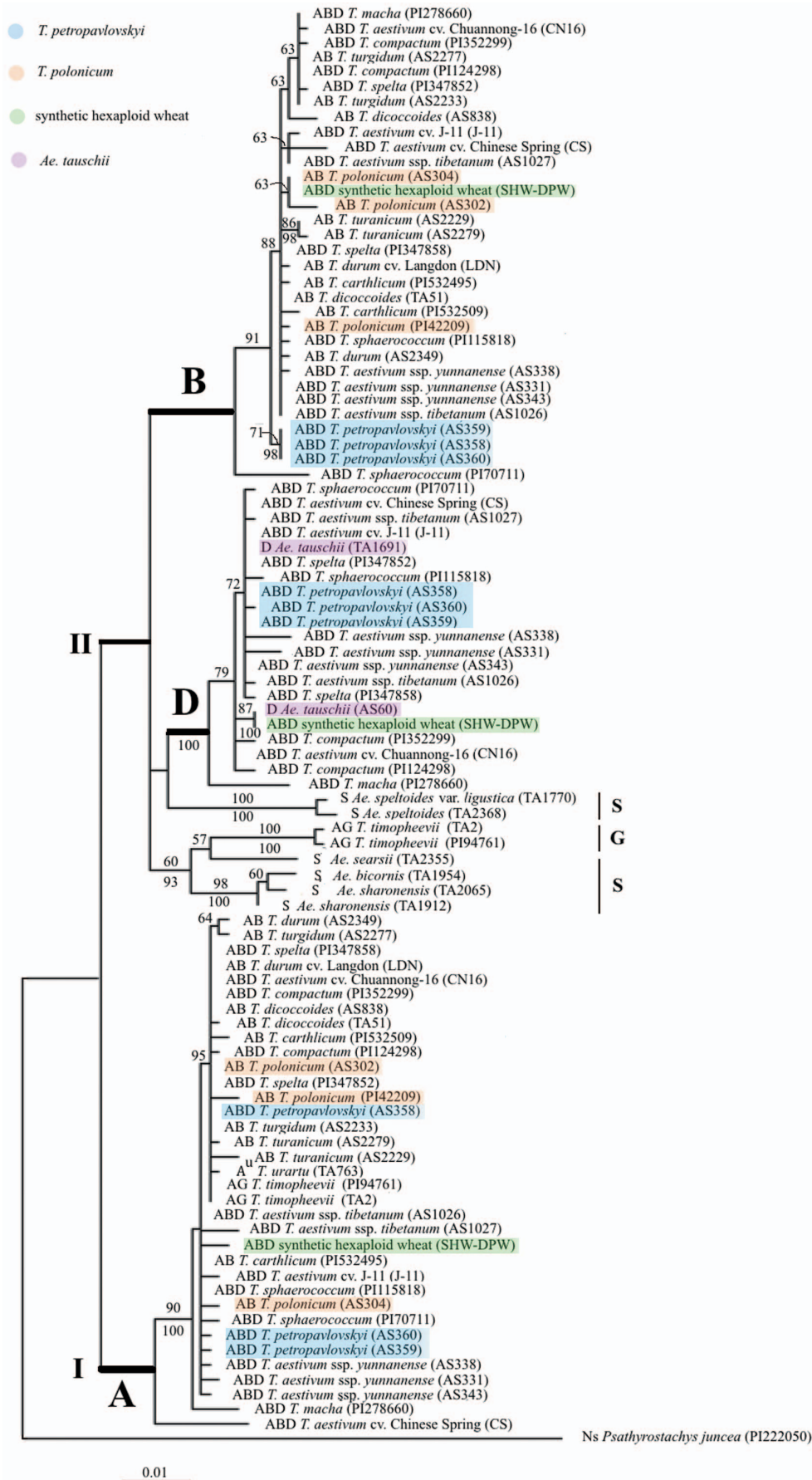


Figure 2. Maximum-likelihood tree from the exon+intron sequences of the *Pgk-1* gene of *T. petropavlovskiy* and its related species. Numbers above nodes are bootstrap values $\geq 50\%$ numbers below nodes are posterior probability values $\geq 90\%$. Genome composition, species name and accession number/cultivar name are indicated for each taxon. doi:10.1371/journal.pone.0071139.g002

created using Maximum likelihood (ML) and Bayesian inference (BI). ML analysis was carried out with PAUP*4.0b10 (Swofford, D. L., Sinauer Associates, <http://www.sinauer.com>), and *Psathyrostachys juncea* (Fischer) Nevski was used as an outgroup. The best-fit models of sequence evolution for ML analysis were estimated using ModelTest v3.0 with Akaike information criteria (AIC) [46]. The optimal models were found to be HKY+G for the exon+intron data, TVM+G for the intron data, and TrN+G for exon data. ML heuristic searches were performed with 100 random addition sequence replications and TBR branch swapping algorithm. The robustness of the trees was estimated using bootstrap support (BS) [47]. ML bootstrapping was performed with 250 replicates, each with three replicates of stepwise random taxon addition, using the same model and parameters. BS-values under 50% were not included in figures.

Bayesian inference (BI) analysis was performed using MrBayes v3.2 [48] with *Psathyrostachys juncea* used as an outgroup. The best-fit models for BI analysis were carried out with AIC using MrModelTest v2.3 (<http://www.ebc.uu.se/systzoo/staff/nylander.html>). The optimal models were found to be GTR+I+G for exon+intron data, GTR+G for intron data and exon data. Four MCMC (Markov Chain Monte Carlo) chains (one cold and three heated) were applied with default setting. In order to make the standard deviation of split frequencies fall below 0.01, 4,200,000 generations for exon+intron data, 3,000,000 generations for intron data, and 5,000,000 for exon data were run. Samples were taken every 100 generations under the best-fit model. For all analyses, the first 25% of samples from each run were discarded as “burn-in” to ensure the stationarity of the chains. Bayesian posterior probability (PP) values were obtained from a majority rule consensus tree generated from the remaining sampled trees. PP-values less than 90% were not included in figures.

Network analysis

The median-joining (MJ) network method [49], and Templeton, Crandall and Sing (TCS) method [36], have been shown to be effective methods for revealing specific progenitor descendant relationships of perennial Triticeae [21,50–52], and were thus performed in this study. Before reconstructing the MJ and TCS networks, a test of recombination was performed using the Phi (Pairwise homoplasy index) method within Splits tree [53]. Building upon this test, the sequences of A ($P=0.64932$), B ($P=0.9999$) and D genomes ($P=1.0$) were used to generate the MJ network, and the exon data ($P=0.7493$) was used to generate the TCS network. MJ network analysis was generated by Network 4.1.1.2 program (Fluxus Technology Ltd, Clare, Suffolk, UK), and TCS haplotype network was performed to evaluate possible genetic relationships between haplotypes with the computer program TCS 1.2.1 [54].

Divergence dating

The potential clock-like evolution of *Pgk-1* sequences was evaluated with a likelihood ratio rate comparing the likelihood scores from the unconstrained and clock-constrained analyses, implemented in PAUP*4.0b10. The molecular clock was rejected

because the substitution rates were significantly heterogeneous ($\chi^2 = 154.90$, $df = 95$, $P = 0.0001$), implying a very poor fit to the molecular clock. Therefore, divergence times with 95% confidence intervals (C.I.) were estimated using the Bayesian relaxed molecular clock method, implemented in BEAST v1.7.1 [55].

The lack of fossils for Triticeae precluded a direct calibration of tree topologies. Instead, molecular dating of the intron data was estimated on the basis of the intron region of the *Pgk-1* gene clock of 0.0051 substitutions per site per MY (million year) [31,38], setting the clock for the divergence of Pooideae and Panicoideae sub-families at 60 MYA [56–58]. Calibration points were performed using a relaxed uncorrelated lognormal molecular clock. MCMC searches were run for 10,000,000 generations under GTR+G model (with the associated parameters specified by MrModelTest as the priors), with the first 2,000,000 discarded as burn-in. Trees were then viewed in FigTree v1.3.1 (<http://tree.bio.ed.ac.uk/>).

Results

Sequence analysis

In all polyploid species considered, the expected number of copies of the *Pgk-1* gene were successfully amplified. The DNA sequence of the *Pgk-1* gene includes 5 exons and 4 introns, which range in length from 1360 bp to 1476 bp, as known from previous study [30,31,38] (Table 2). The lengths of exon+intron, exon and intron data sets were 1466, 894 and 572 bp, respectively. As expected, the level of nucleotide variation in the exon region (88 variable sites and 36 parsimony-informative sites) was lower than that in the intron region (107 variable sites and 80 parsimony-informative sites). The average content of G+C of exon+intron, and exon was 42.9, and 47.3%, respectively, and the transition/transversion ratio was 2.13, and 2.24, respectively. The alignment of the exon sequence was unambiguous and without gaps. Gaps were, on the contrary, present in introns. In particular, apart from single nucleotide substitutions and deletions, three significant indels (insertion/deletion) (indel 1, 2, 3) were found (Fig. 1). Indel 1 was located at position 67–72 of the A genome; indel 2 mapped at position 563–568 and had a 6 bp deletion specific for A genome. Unexpectedly, the *T. aestivum* ssp. *yunnanense* (AS338) did not show the indel 2. Indel 3 was presented at position 1220–1308 and had a 89 bp insertion in the G genome of *T. timopheevii*.

Phylogenetic analyses

Using *Psathyrostachys juncea* as an outgroup, the three data sets corresponding to exon+intron, exon and intron were used phylogenetic analyses (ML and BI) were carried out. ML analysis of the exon+intron data generated a single phylogenetic tree ($-\ln\text{likelihood} = 4092.85$), with the following parameters: $A = 0.26$; $C = 0.21$; $G = 0.23$, $T = 0.30$, gamma shape parameter = 0.29. Bayesian analysis of the same data recovered the same topology. In Figure 2, the ML tree is reported with values of the bootstrap support (BS) above and posterior probabilities (PP) below branches.

The ML tree of Figure 2 indicates that all homoeologous *Pgk-1* sequences from polyploid accessions are grouped with those of the

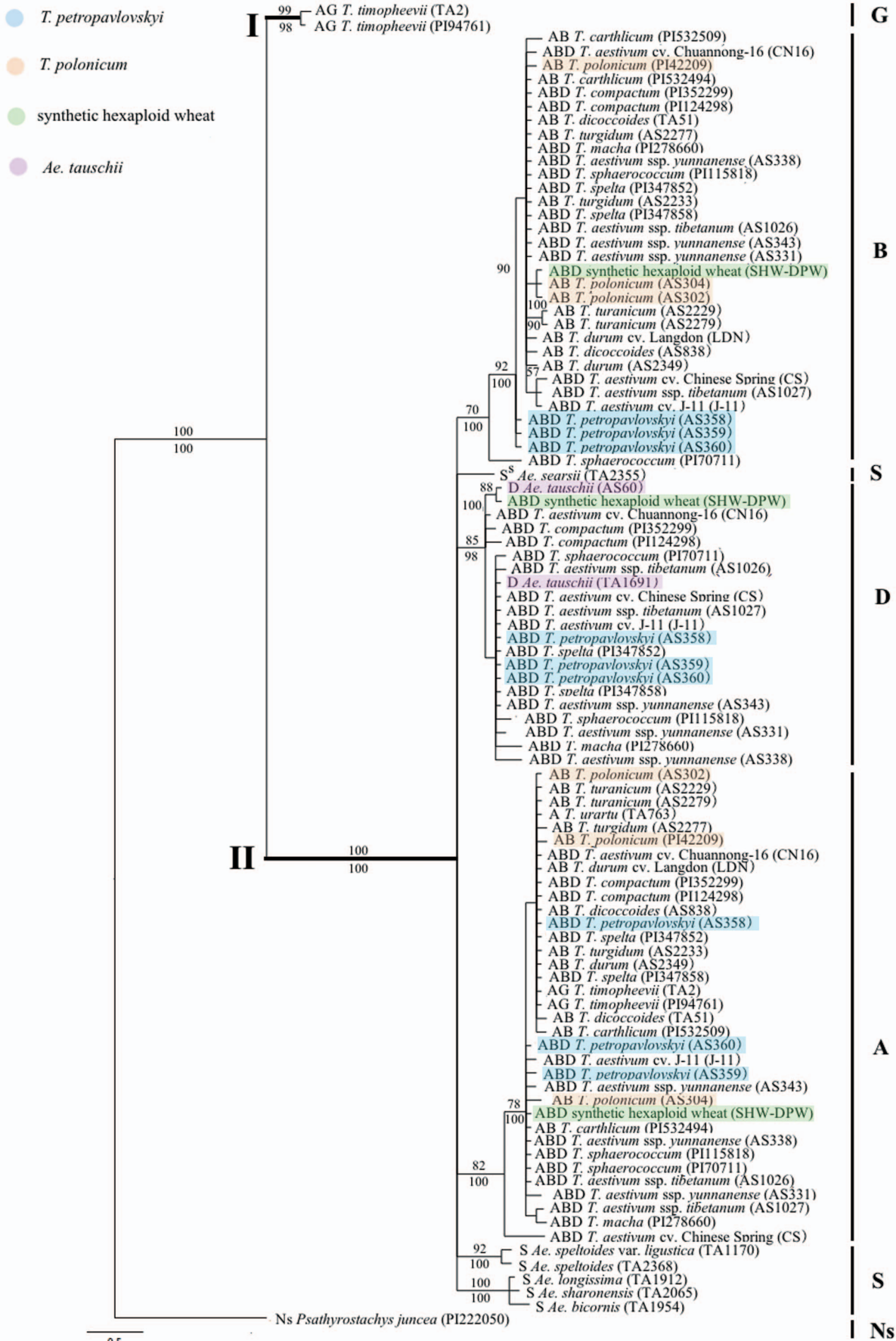


Figure 3. Bayesian tree inferred from the intron sequences of the *Pgk-1* gene of *T. petropavlovskiyi* and its related species. Numbers above nodes are bootstrap values >50%; below nodes are Bayesian posterior probability values >90%. Genome composition, species name and accession number/cultivar name are indicated for each taxon. doi:10.1371/journal.pone.0071139.g003

diploid parental species. The tree has two major clades: the one including sequences of the A genome and the second those derived from the genomes B, D, G and S. In Clade I, the A genome specific sequences from three *T. petropavlovskiyi* accessions and from *Triticum* species (except *T. aestivum* cv. Chinese Spring) formed a group with 90% bootstrap value and 100% posterior probabilities support. One *T. petropavlovskiyi* accession (AS358), together with two accessions of *T. polonicum* (AS304 and PI42209), formed a subclade, with bootstrap value of 95%. In Clade II, the B genome sequences from three *T. petropavlovskiyi* accessions clustered together in a well supported (71% BS and 98% PP) subclade. SHW-DPW had a topology contiguous with two accessions of *T. polonicum* (AS302 and AS304), with 63% bootstrap support. The sequences from the D genomes mapped to two subclades. One consisted of three accessions of *T. petropavlovskiyi*, eleven of *T. aestivum* and one of *Ae. tauschii* (TA1691), with 72% bootstrap support. The second one included SHW-DPW and *Ae. Tauschii* (AS60) with 87% BS and 100% PP.

ML analysis of the intron data yielded a single phylogenetic tree ($-\ln$ likelihood = 1804.87), with parameters: A = 0.29; C = 0.18; G = 0.17; T = 0.36 and gamma shape parameter = 0.60. The Bayesian analysis generated the same topology, as illustrated in Figure 3. Two major clades are evident: Clade I includes only the *Pgk-1* sequences from the G genome with a high bootstrap support (99% BS, 98% PP). The Clade II includes A, B, D and S genomes sequences and is congruent with the tree inferred from the exon+intron data, except for nodes presenting different statistical support. In Clade II, sequences from the B genome clustered together with a good support (70% BS and 100% PP). In this B genome clade, with the exception of *T. sphaerococcum* (PI70711), all accessions were grouped in a well supported subclade (92% BS and 100% PP); the three accessions of *T. petropavlovskiyi* were separated from other *Triticum*. In the A genome subclade, *T. petropavlovskiyi* (AS359) mapped together with *T. polonicum* (AS304) (78% BS and 100% PP).

Exon data generated a single ML phylogenetic tree ($-\ln$ likelihood = 2150.29), with the following parameters A = 0.24, C = 0.21, G = 0.27, T = 0.28, gamma shape parameter = 0.38. ML and BI analysis of the same data supported a similar topology. Figure 4 reports the ML tree of exon sequences which includes two major clades: Clade I, consisting of A genome sequences: *T. carthlicum* (PI532509) and *T. petropavlovskiyi* (AS360) are mapped in the same group (64% BS and 95% PP). In the second clade (Clade II), which includes B, D, G, and S genomes, the B genome sequences of three accessions of *T. petropavlovskiyi* grouped together (64% BS and 93% PP), separated from other accessions. *Ae. tauschii* (TA1691) and *Ae. speltoides* (TA2368) clustered in a subclade (76% BS and 97% PP).

Network analysis

To highlight the relationships among haplotypes of the *Pgk-1* sequence, network methods were employed and the exon+intron data (Fig. 5) and exon data (Fig. 6) were considered. In Figure 5, each circular network node represents a haplotype, with node size being proportional to number of its isolates. Mv (median vectors representing missing intermediates) shows unsampled nodes

inferred from the MJ network analysis. The number on the branches indicates the positions of the mutations. Network loops represent either true reticulation events or alternative genealogies in closely related lineages. MJ analysis of the *Pgk-1* exon+intron data recovered groupings corresponding to clades revealed by ML phylogeny. *T. petropavlovskiyi*, as expected, was present in three clusters (A, B and D), representing the A, B and D genomes. Most accessions of *T. aestivum*, except *T. aestivum* cv. Chinese spring, were included in the A-type. The A-type sequences of three accessions of *T. petropavlovskiyi* were included in subgroups I and II. *T. petropavlovskiyi* (AS358) and *T. polonicum* (AS302) were placed at a central branching point. Meanwhile, in the B-type cluster, three accessions of *T. petropavlovskiyi* grouped together in subgroup III, and *T. polonicum* (AS304) together with SHW-DPW in subgroup IV. In the D-type cluster, *T. petropavlovskiyi* (AS358 and AS359), *T. aestivum* cv. Chinese Spring, *T. aestivum* cv. J-11, *T. aestivum* ssp. *yunnanense* (AS343) and *T. spelta* (PI347852) resulted included in subgroup V, while the sequences from the amphiploid SHW-DPW and *Ae. tauschii* formed the distinct subgroup VI. The TCS procedure [36] was used to illustrate haplotype relationships among accessions. TCS defined a 95% parsimony connection limit of 13 steps for exon alignment of fifty haplotypes derived from 96 sequences (Fig. 6). The TCS network consisted of three major haplotypic groups corresponding to the A, B and D genomes. The length of the branches between two nodes was proportional to the nucleotidic difference. In TCS analysis, *T. petropavlovskiyi* (AS360) shows a close haplotype relationship with *T. carthlicum* from Xinjiang, China, supporting the exon results of the ML and BI analyses. Two further differences between the TCS and MJ were noted. Firstly, the haplotype of the D genome of *T. macha* (PI278660) appeared related to haplotypes of the B genome accessions. Secondly, *Ae. tauschii* (TA1691) showed a close haplotype relationship with the S genome of *Ae. speltoides* var. *ligustica*.

Molecular dating

The BEAST analysis of the intron region of *Pgk-1* was used to derive a time-calibrated phylogenetic tree (Fig. 7). Under a lognormal relaxed clock, rate variation was equal to 0.96 (95% C.I., 0.67–1.39), supporting the adoption of the relaxed clock method. The Yule prior was equal to 0.47 (95% C.I., 0.37–0.63) and five homoeologous types of the *Pgk-1* gene, A-, B-, D, G- and S-type, clustered in distinct clades. The divergence time of the A, B, and D genomes of *T. petropavlovskiyi* was estimated equal to 1.13 (95% C.I., 0.65–1.75), 1.02 (95% C.I., 0.24–1.51), and 0.73 MYA (95% C.I., 0.41–1.01), respectively. The split between *Pgk-1* A and D genomes of *T. petropavlovskiyi* and its putative diploid genome donor, *T. polonicum* and *Ae. tauschii*, took place around 0.74–1.13 and 0.33–0.73 MYA, respectively. The B genome diverged from the S genome at 2.27 MYA (95% C.I., 1.68–3.19), while the divergence time of *T. petropavlovskiyi* and *T. polonicum* was 0.68–0.91 MYA for A genome. The divergence time of *T. petropavlovskiyi* and *T. polonicum* and the B genome was 0.34–0.78 MYA. The divergence time of *T. petropavlovskiyi* from hexaploid wheat resulted equal to 0.14–0.33 MYA (A genome), 0.16–0.69 MYA (B genome) and 0.11–0.27 MYA (D genome), respectively (node 1–node 9). In

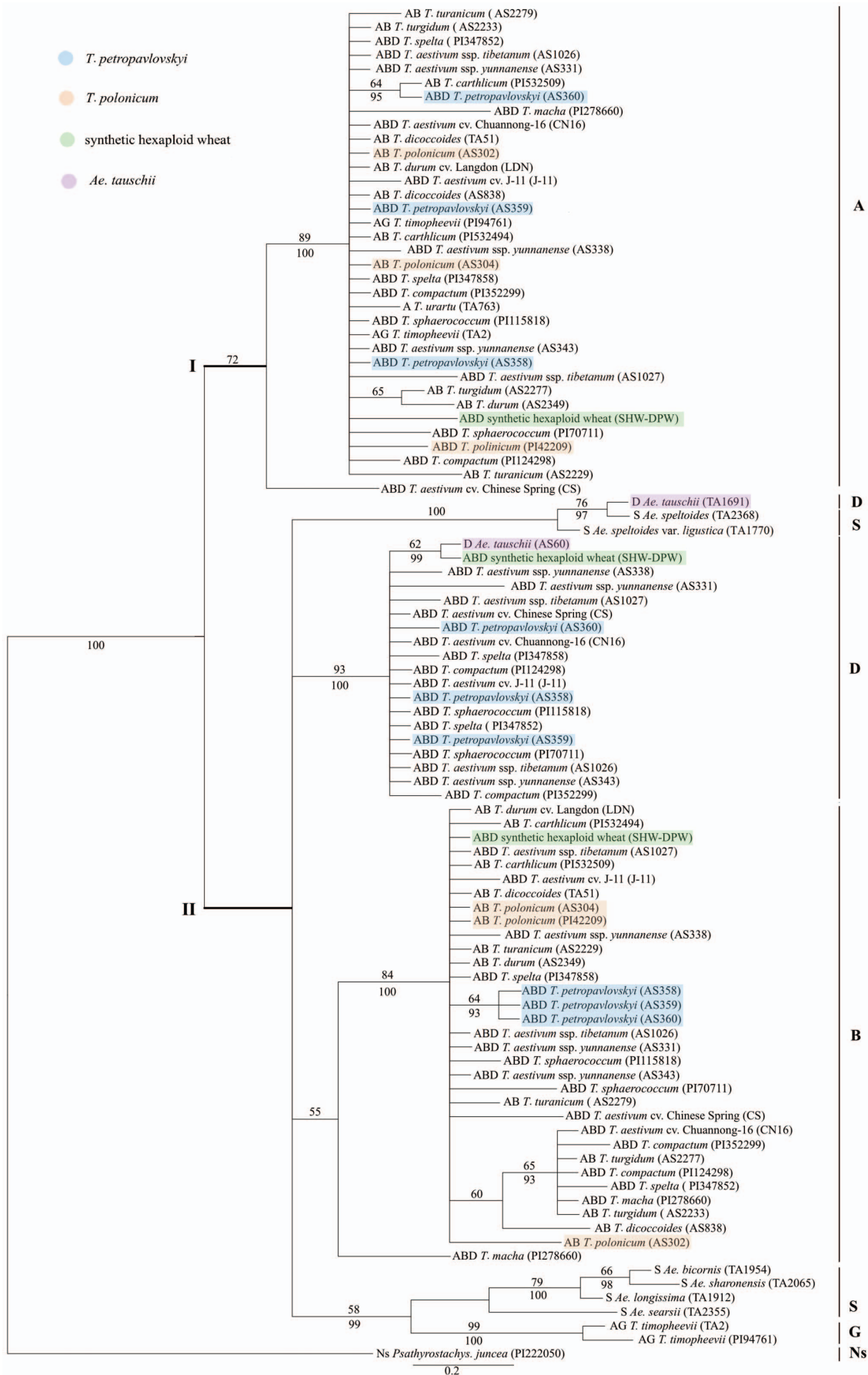


Figure 4. Maximum-likelihood tree inferred from the exon sequences of the *Pgk-1* gene of *T. petropavlovskiyi* and its related species. Numbers above nodes are bootstrap values $\geq 50\%$; below nodes are Bayesian posterior probability values $\geq 90\%$. Genome composition, species name and accession number/cultivar name are indicated for each taxon. doi:10.1371/journal.pone.0071139.g004

the tree, the tetraploid wheat *T. polonicum* diverged earlier than *T. petropavlovskiyi*. In the A genome, the divergence time of *T. petropavlovskiyi* (AS358) was later than the other two accessions of *T. petropavlovskiyi*. On the contrary, in B and D genomes, the divergence time was earlier than the other two accessions. Additionally, the divergence time of B genome of *T. petropavlovskiyi* was earlier than other three Chinese endemic wheat landraces. Between *T. petropavlovskiyi* and SHW-DPW, a significant difference was observed: in the genomes A, B and D, the divergence time of SHW-DPW was later than *T. petropavlovskiyi*.

Discussion

Relationships between *T. petropavlovskiyi* and hexaploid wheat taxa

Based on cytological results, Yao et al. [3] and Chen et al. [23] suggested that the B genome was responsible for the difference between *T. petropavlovskiyi* and *T. aestivum* cv. Chinese Spring, and that two pairs of chromosomes, one identified as chromosome 6B [2], were involved. The allelic variation at the HMW glutenin subunits loci, *Gli-1* and *Gli-2*, supported the cytological results [59]. Also, Yang et al. [10] reported that *T. petropavlovskiyi* differed from *T. spelta* in at least one or two pairs of chromosomes. Results based on molecular markers, including A-PAGE, SDS-PAGE, STS-PAGE, SSR and RFLP, indicate that *T. petropavlovskiyi* is genetically distinct from other Chinese endemic wheat landraces [5,59].

Our ML and BI study of the *Pgk-1* gene indicates that the A and D genomes of *T. petropavlovskiyi* are basically shared with *T. spelta*, *T. compactum* and *T. sphaerococcum*. When the B genome is considered, *T. petropavlovskiyi* groups in one subclade, comparatively distantly related to *T. aestivum*. In addition, the B-type of MJ network shows that the accessions of *T. petropavlovskiyi* (subgroup III) are distinct from those of other species. SHW-DPW, a synthetic hexaploid wheat with both genomes of *T. polonicum* and *Ae. tauschii* [26], based on the *Pgk-1* gene is characterized by the A, B and D genomes, the SHW-DPW is distant from those of *T. petropavlovskiyi* in B and D genomes.

Relationships between *T. petropavlovskiyi* and the tetraploid wheats

Morphologically, the spikelet of *T. petropavlovskiyi* are similar to those of *T. turanicum* and *T. polonicum* [7,26]. Moreover, the cytology of interspecific hybrids between *T. petropavlovskiyi* and tetraploid wheats support a closer relationship with the AABB genomes, compared to wheats with the AAGG genome [2]. According to Akond and Watanabe [24], *T. petropavlovskiyi* is more closely related to *T. polonicum* than to *T. durum* or *T. turgidum*. However, Arbutov et al. [60] and Efremova et al. [18] report that the genes supporting the elongated glumes in *T. polonicum* and *T. petropavlovskiyi* are not allelic.

In the present study, based on the ML and BI analyses of the genome A *Pgk-1* gene, two accessions of *T. polonicum* have a common topology with *T. petropavlovskiyi*, while the TCS analysis of exon data indicates that the haplotype of A genome in *T. petropavlovskiyi* (AS360) is more closely related to *T. carthilicum* than

to other wheats. The phylogenetic analyses and MJ network specific for the B genome shows that three accessions of *T. petropavlovskiyi* group together, and are topologically distant from those of tetraploid wheats. This finding indicates that the B genome of *T. petropavlovskiyi* diverge from the one of tetraploid species. Yao et al. [3] and Chen et al. [23] also recognized cytologically that the B genome of *T. petropavlovskiyi* was different from those of hexaploid wheats.

Relationships between *T. petropavlovskiyi* and *Ae. tauschii*

Based on RFLPs, Ward et al. [5] found that *T. petropavlovskiyi* is genetically more closely related to accessions of *Ae. tauschii* from Iran than from China. Yang et al. [10] concluded that *T. petropavlovskiyi* is derived from a hybrid between *Ae. tauschii* and a presumed *T. polonicum* genotype. However, the phylogenetic analysis of the *Acc-1* genes indicates that the D genome of *T. petropavlovskiyi* is very similar to the D genome orthologs of *T. aestivum* and only distantly related to *Ae. tauschii* [26]. In the present study, D genomes of *Ae. tauschii* belong to two different clusters. One groups with SHW-DPW, *T. compactum* and *T. aestivum* cv. Chuannong-16, while *T. petropavlovskiyi*, *T. macha*, *T. spelta*, *T. sphaerococcum* and three Chinese endemic wheats are included in a clade with *Ae. tauschii* (TA1691). Based on TCS analysis, *T. petropavlovskiyi* shares common topologies with the hexaploid species D genomes. The sequence of the *Pgk-1* gene from TA1691 is significantly different from those of other accessions of *Ae. tauschii*, in agreement with Huang et al. [31]. Together, available results supports that the D genome of *T. petropavlovskiyi* is similar to D genome orthologs of *T. aestivum* and only distantly related to *Ae. tauschii*.

The divergence time of the *T. petropavlovskiyi*

Huang et al. [31] report that the diploid progenitors of the A, B and D genomes present in diploid, tetraploid, and hexaploid wheats radiated between 2.5 and 4.5 MYA. We report that the divergence time of the A, B and D genomes corresponds to 2.61 (95% C.I., 1.77–3.55), 2.27 (95% C.I., 1.69–3.19) and 2.05 MYA (95% C.I., 1.40–2.81), respectively. The divergence of the B genome of *T. petropavlovskiyi* from those of other wheats is dated in this paper is from 0.16 (95% C.I., 0–0.38) to 0.69 MYA (95% C.I., 0.55–0.70), the earliest date for the four Chinese endemic wheat landraces we considered.

Concerning the A and D genome of *T. petropavlovskiyi*, the resulting divergence times are 0.14–0.33 and 0.11–0.27 MYA, respectively, values similar to those of other hexaploid species. The divergence time of A genome of *T. petropavlovskiyi* from *T. polonicum* varies from 0.68 (95% C.I., 0.41–0.71) to 0.90 MYA (95% C.I., 0.45–1.41), a divergence earlier than those between *T. petropavlovskiyi* and hexaploid species. The divergence time results indicate that *T. polonicum* may have played a role in the evolutionary history of *T. petropavlovskiyi*.

The possible origin of *T. petropavlovskiyi*

This study shows that the *Pgk-1* sequences of the A genome of *T. petropavlovskiyi* group with *T. polonicum*. For the *Pgk-1* locus of the D genome, the accessions of *Ae. tauschii* just cluster with the

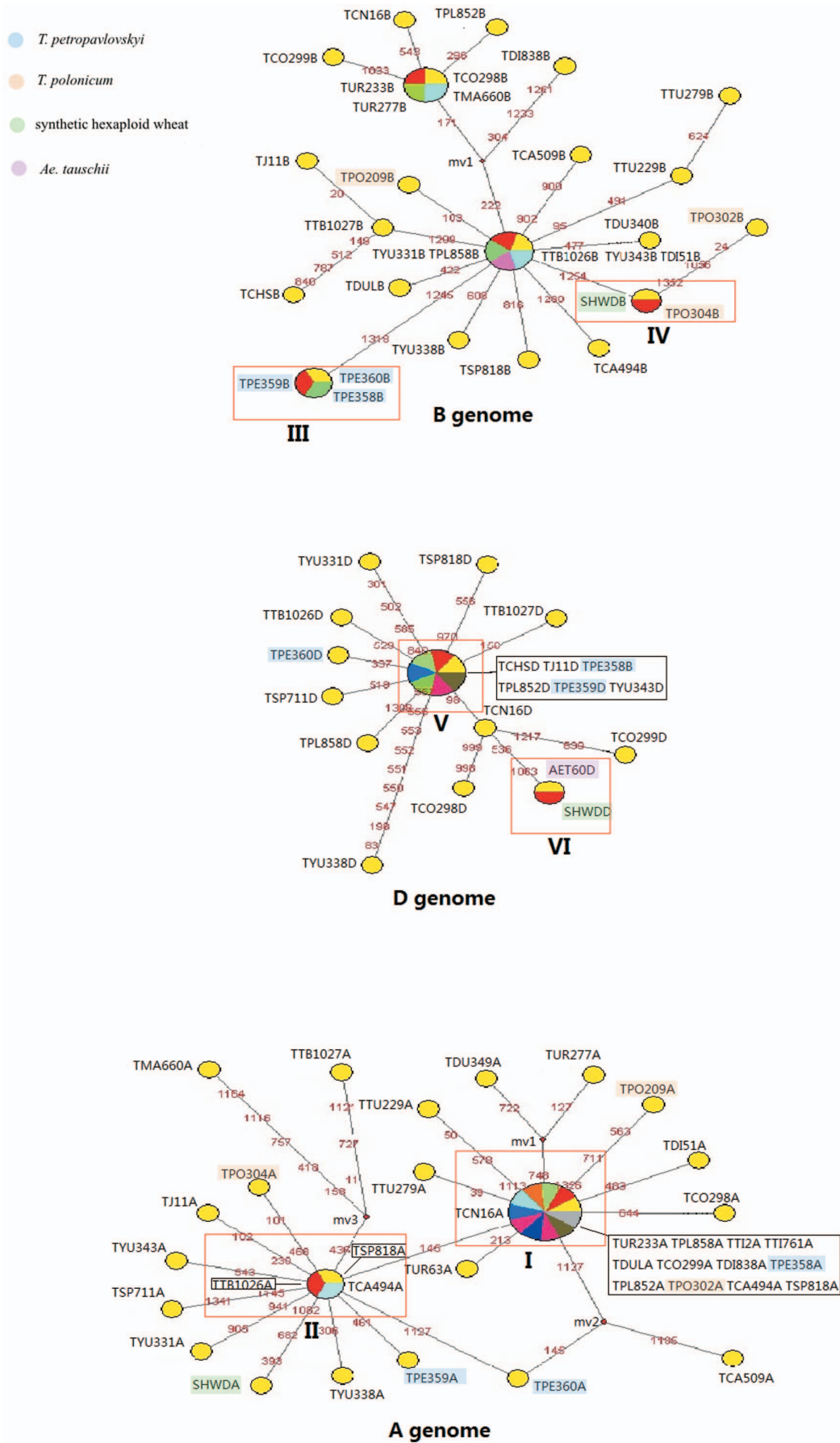


Figure 5. Median-joining (MJ) network inferred from the exon+intron sequences of the *Pgk-1* gene of *T. petropavlovskyi* and its related species. Abbreviations of the species names in the MJ network are listed in Table 1. Haplotypes in the network are represented by circles. Distance between nodes is proportional to the number of nucleotide substitutions among sequences. doi:10.1371/journal.pone.0071139.g005

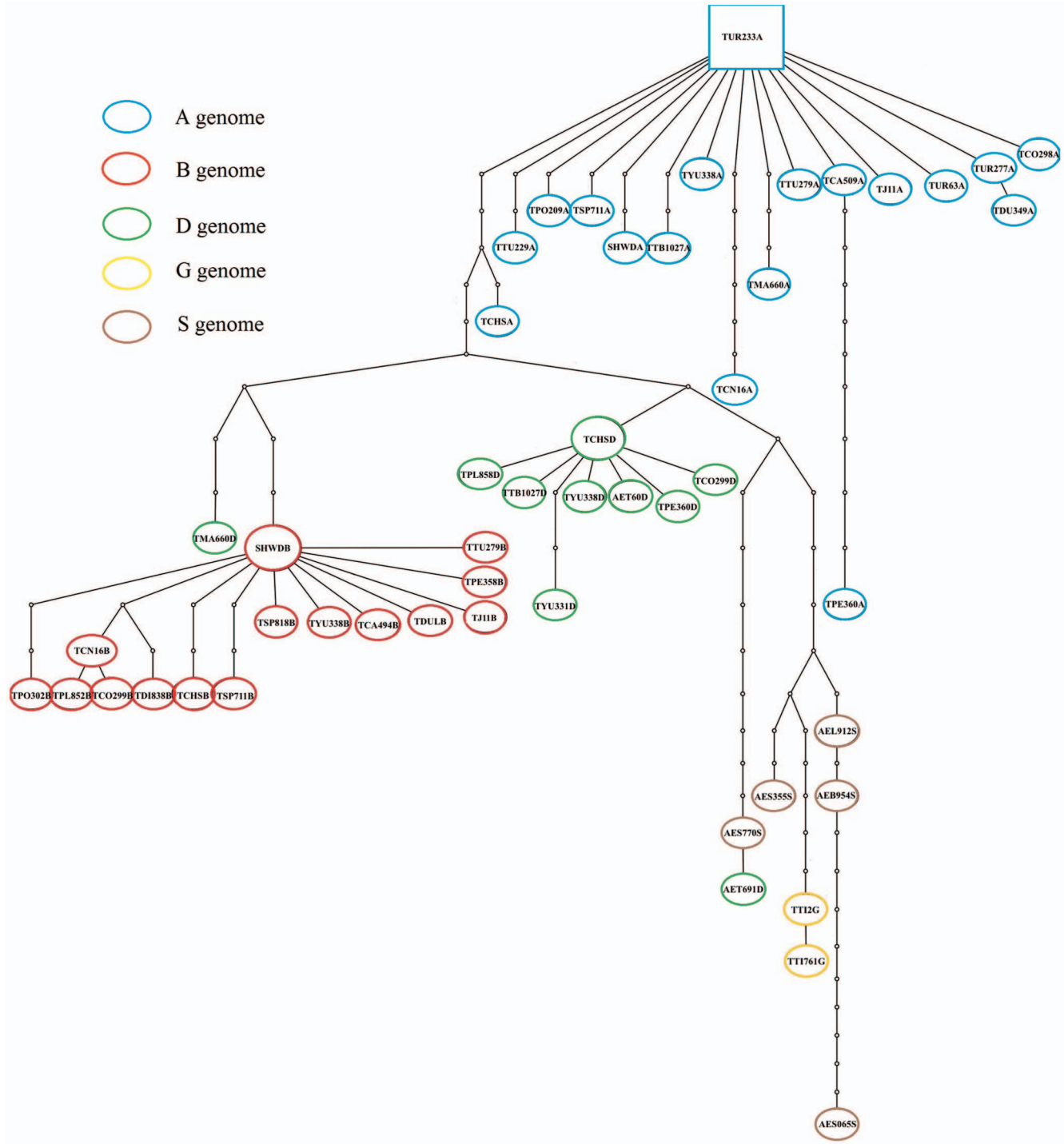


Figure 6. TCS network inferred from the exon sequences of the *Pgk-1* gene of *T. petropavlovskyi* and its related species. Abbreviations of species names are listed in Table 1. Haplotypes in the network are represented by circles of different color corresponding to the genomes indicated. doi:10.1371/journal.pone.0071139.g006

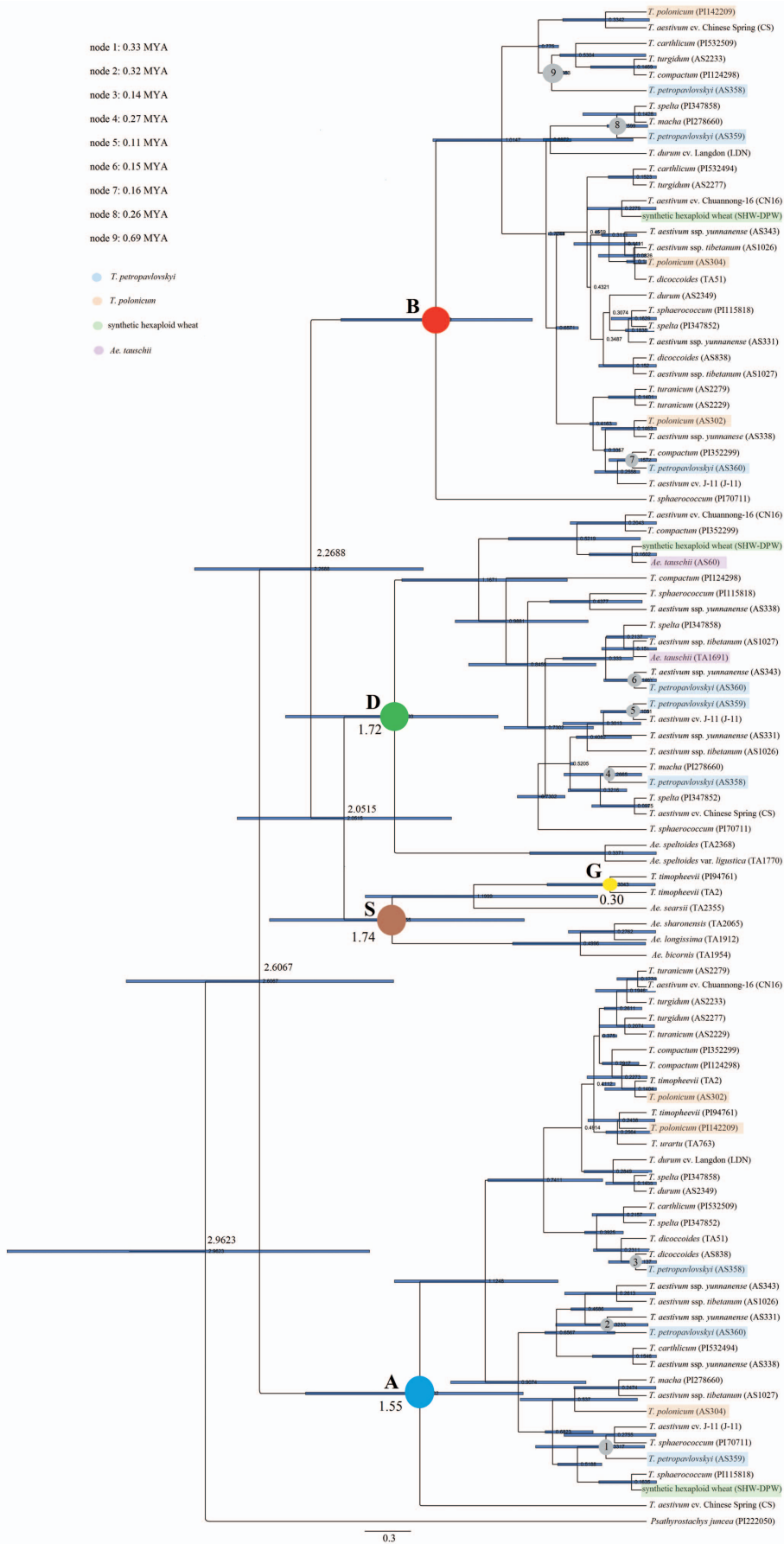


Figure 7. Time-calibrated tree based on intron region of the *Pgk-1* sequence of *T. petropavlovskiyi* and related species using a Bayesian relaxed BEAST clock method. The different color of node labeled the genome information of the subclade. Numbers at nodes provide the estimated divergence dates.
doi:10.1371/journal.pone.0071139.g007

amphiploid SHW-DPW. The *Pgk-1* B genome data indicate that *T. petropavlovskiyi* is distantly related to the other three Chinese endemic wheat landraces. The MJ network results are congruent with the results reported above. Also the TCS analysis supports the conclusion that the relationship among haplotypes of *T. petropavlovskiyi* and *T. polonicum* have very similar A and B genomes. We report a distant relationship between *T. petropavlovskiyi* and *Ae. tauschii*. We conclude that *T. petropavlovskiyi* is neither derived from an independent allopolyploidization event nor from a single mutation in *T. aestivum*. It is most likely that *T. petropavlovskiyi* has an origin starting with a natural cross between *T. aestivum* and *T. polonicum*, with that event taking place around 0.78 MYA.

References

- Yen C, Yang JL, Luo MC (1988) The origin of the Tibetan weedrace of hexaploid wheat, Chinese Spring, Chengdu guangtuo and other landraces of white wheat complex from china. In Proceedings of the 7th International Wheat Genetics Symposium Miller TE, Koebner RMD, eds. Cambridge, UK. pp 175–179.
- Chen Q, Sun YZ, Dong YS (1985) Cytogenetical studies on interspecific hybrids of Xinjiang wheat. Acta Agron Sin 11: 23–28.
- Yao JX (1983) Research on a new species in *Triticum*-Xinjiang wheat with rice-like spike. Hereditas (Beijing) 5: 17–20.
- Kim HS, Ward RW (2000) Pattern of RFLP-based genetic diversity in germplasm pools of common wheat with different geographical or breeding program origins. Euphytica 115: 197–208.
- Ward RW, Yang ZL, Kim HS, Yen C (1998) Comparative analyses of RFLP diversity in landraces of *Triticum aestivum* and collections of *Aegilops tauschii* from China and South Asia. Theor Appl Genet 96: 312–318.
- Jakubtsiner MM (1959) K poznaniyu pshenits Kitaja/A contirbution to the knowledge of the wheats of China. Bot J 44: 1425–1436.
- Udaczin RA, Miguschova EF (1970) Novoe v poznanii roda *Triticum* L. Venstnik S-Kh Nauki 9: 20–24.
- Riley R, Coucolli H, Chapman V (1967) Chromosomal interchanges and the phylogeny of wheat. Heredity 22: 233–247.
- Shao QQ, Li CS, Basang CR (1980) Semi-wild wheat from Xizang (Tibet). Acta Genet Sin 7: 150–156.
- Yang WY, Yen C, Yang JL (1992) Cytogenetic study on the origin of some special Chinese landraces of common wheat. Wheat Inform Serv 75: 14–20.
- Dorofeev VF, Filatenko AA, Migushova EF, Udaczin RA, Jakubziner MM (1979) Flora of Cultivated plants. In Wheat Dorofeev VF, Migushova EKolos, L., eds. pp 1–384.
- Watanabe N, Imamura I (2002) The inheritance and chromosomal location of a gene for long glume phenotype in *Triticum petropavlovskiyi* Udacz. et Migusch. J Genet Breed 57: 221–227.
- Watanabe N, Sekiya T, Sugiyama K, Yamagishi Y, Imamura I (2002) Telosomic mapping of the homoeologous genes for the long glume phenotype in tetraploid wheat. Euphytica 128: 129–134.
- Wang HY, Huang XQ, Röder MS, Börner A (2002) Genetic mapping of loci determining long glumes in the genus *Triticum*. Euphytica 123: 287–293.
- Yen C, Yang JL, Liu XD, Li LR (1983) The distribution of *Aegilops tauschii* Cosson in China with reference of the origin of the Chinese common wheat. In Proceedings of the 7th International Wheat Genetics Symposium, Sakamoto S, eds. Kyoto, Japan. pp 55–58.
- Yang RW, Zhou YH, Zheng YL (2001) Analysis on chromosome C-banding of dwarf polish wheat (*Triticum polonicum*). J Sichuan Agric Univ 19: 112–114.
- Liu GX, Zhou YH, Zheng YL, Yang RW, Ding CB (2002) Morphological and cytological studies of dwarfing polish wheat (*Triticum turgidum* conv. *polonicum*) from Xinjiang, China. J Sichuan Agric Univ 20: 189–193.
- Efremova TT, Maystrenko OI, Laikova LI, Arbutzova VS, Popova OM (2000) Comparative genetic analysis of hexaploid wheats *Triticum petropavlovskiyi* Udacz. et Migusch. and *Triticum aestivum* L. Russ J Genet 36: 1142–1148.
- Yang RW, Zhou YH, Zheng YL, Hu C (2000) Genetic differences and the relationship of gliadin between *Triticum polonicum* and *Triticum petropavlovskiyi*. J Triticeae Crops 20: 1–5.
- Chen QF (1999) Discussion on origin of Chinese endemic wheat. Guizhou Agri Sci 27: 20–25.
- Goncharov NP (2005) Comparative genetic analysis—a base for wheat taxonomy revision. Czech J Genet Plant Breed 41: 52–55.
- Akond ASGM, Watanabe N, Furuta Y (2008) Comparative genetic diversity of *Triticum aestivum*-*Triticum polonicum* introgression lines with long glume and *Triticum petropavlovskiyi* by AFLP-based assessment. Genet Resour Crop Evol 55: 133–141.
- Chen PD, Liu DJ, Pei GZ, Qi LL, Huang L (1988) The chromosome constitution of three endemic hexaploid wheats in western China. In Proceedings of the 7th International Wheat Genetics Symposium Miller TE, Koebner RMD, eds. Cambridge, UK. pp 75–80.
- Akond ASGM, Watanabe N (2005) Genetic variation among Portuguese landraces of ‘Arrancada’ wheat and *Triticum petropavlovskiyi* by AFLP-based assessment. Genet Resour Crop Evol 52: 619–628.
- Kang HY, Wang Y, Yuan HJ, Jiang Y, Zhou YH (2008) A new synthesized 6x-wheats, derived from dwarfing polish wheat (*Triticum polonicum* L.) and *Aegilops tauschii* Cosson. Intern J Agric Res 3: 252–260.
- Kang HY, Fan X, Zhang HQ, Sha LN, Sun GL, et al. (2010) The origin of *Triticum petropavlovskiyi* Udacz. et Migusch.: demonstration of the utility of the genes encoding plastid acetyl-CoA carboxylase sequence. Mol Breed 25: 381–395.
- Sang T (2002) Utility of low-copy nuclear gene sequences in plant phylogenetics. Crit Rev Biochem Mol Biol 37: 121–147.
- Smith J, Funke M, Woo V (2006) A duplication of *gycx* predates divergence within tribe Coronantherae (Gesneriaceae): phylogenetic analysis and evolution. Plant Syst Evol 261: 245–256.
- Chalupska D, Lec HY, Farris JD, Evrard A, Chalhou B, et al. (2008) *Ace* homoeoloci and the evolution of wheat genomes. Proc Natl Acad Sci 105: 9691–9696.
- Huang SX, Sirikhachornkit A, Farris JD, Su XJ, Gill BS, et al. (2002) Phylogenetic analysis of the acetyl-CoA carboxylase and 3-phosphoglycerate kinase loci in wheat and other grasses. Plant Mol Biol 48: 805–820.
- Huang SX, Sirikhachornkit A, Su XJ, Farris J, Gill BS, et al. (2002) 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 99: 8133–8138.
- Golovnina KA, Glushkov SA, Blinov AG, Mayorov VI, Adkison LR, et al. (2007) Molecular phylogeny of the genus *Triticum*. Plant Syst Evol 264: 195–216.
- Yan C, Sun GL, Sun DF (2011) Distinct origin of the Y and St genome in *Elymus* species: evidence from the analysis of a large sample of St genome species using two nuclear genes. PLoS ONE 6: e26853.
- Wu F, Mueller LA, Crouzillat D, Pétiard V, Tanksley SD (2006) Combining bioinformatics and phylogenetics to identify large sets of single-copy orthologous genes (COSII) for comparative, evolutionary and systematic studies: a test case in the Euastrid plant clade. Genetics 174: 1407–1420.
- Sun GL, Ni Y, Daley T (2008) Molecular phylogeny of *RPB2* gene reveals multiple origin, geographic differentiation of H genome, and the relationship of the Y genome to other genomes in *Elymus* species. Mol Phylogenet Evol 46: 897–907.
- Sun GL, Salomon B (2009) Molecular evolution and origin of tetraploid *Elymus* species. Breed Sci 59: 487–491.
- 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*). Mol Phylogenet Evol 39: 70–82.
- Fan X, Sha LN, Zeng J, Kang HY, Zhang HQ, et al. (2012) Evolutionary dynamics of the *Pgk-1* gene in the polyploid genus *Kengyilia* (Triticeae: Poaceae) and its diploid relatives. PLoS ONE 7: e31122.
- Gornicki P, Farris J, King I, Podkowinski J, Gill BS, et al. (1997) Plastid-localized acetyl-CoA carboxylase of bread wheat is encoded by a single gene on each of the three ancestral chromosome sets. Proc Natl Acad Sci 94: 14179–14184.
- Kilian B, Özkan H, Deusch O, Effgen S, Brandolini A, et al. (2007) Independent wheat B and G genome origins in outcrossing *Aegilops* progenitor haplotypes. Mol Biol Evol 24: 217–227.

Acknowledgments

We thank Dr. Norman Ellstrand, University of California Riverside, for his help in English polishing. We also thank two anonymous reviewers for their very useful comments on this manuscript.

Author Contributions

Conceived and designed the experiments: QC HYK YHZ. Performed the experiments: QC YW MYZ. Analyzed the data: XF LNS JZ. Contributed reagents/materials/analysis tools: HQZ RWY LZ CBD LLX. Wrote the paper: QC HYK YHZ.

41. Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* 12: 13–15.
42. Doyle JJ, Doyle JL (1999) Nuclear protein-coding genes in phylogeny reconstruction and homology assessment: some examples from Leguminosae. In *The Molecular Systematics and Plant Evolution* Hollingsworth PM, Bateman RM, Gornall RJ, eds. Taylor and Francis, London. pp 229–254.
43. Fan X, Sha LN, Yang RW, Zhang HQ, Kang HY, et al. (2009) Phylogeny and evolutionary history of *Lymus* (Triticeae; Poaceae) based on a single-copy nuclear gene encoding plastid acetyl-CoA carboxylase. *BMC Evol Biol* 9: 247.
44. Thompson JD, Plewniak F, Poch O (1999) A comprehensive comparison of multiple sequence alignment programs. *Nucleic Acids Res* 27: 2682–2690.
45. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731–2739.
46. Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics* 14: 817–818.
47. Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791.
48. Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754–755.
49. Allaby RG, Brown TA (2001) Network analysis provides insights into evolution of 5S rDNA arrays in *Triticum* and *Aegilops*. *Genetics* 157: 1331–1341.
50. Wang HY, Wang XE, Chen PD, Liu DJ (2007) Assessment of genetic diversity of Yunnan, Tibetan, and Xinjiang wheat using SSR markers. *J Genet Genomics* 34: 623–633.
51. Bordbar F, Rahiminejad MR, Saicidi H, Blattner FR (2011) Phylogeny and genetic diversity of D-genome species of *Aegilops* and *Triticum* (Triticeae, Poaceae) from Iran based on microsatellites, ITS, and *trnL-F*. *Plant. Syst Evol* 291: 117–131.
52. Yan C, Sun GL (2011) Nucleotide divergence and genetic relationships of *Pseudoroegneria* species. *Biochem Syst Ecol* 39: 309–319.
53. Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* 23: 254–267.
54. Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Mol Ecol* 9: 1657–1659.
55. Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 29: 1969–1973.
56. Gaut BS (1998) Molecular clocks and nucleotide substitution rates in higher plants. In *The Evolutionary Biology* Hecht MK, ed. Plenum Press New York, pp 93–120.
57. Wolfe KH, Gouy M, Yang YW, Sharp PM, Li WH (1989) Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. *Proc Natl Acad Sci* 86: 6201–6205.
58. Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc Natl Acad Sci* 84: 9054–9058.
59. Wei YM, Zheng YL, Liu DC, Zhou YH, Lan XJ (2002) HMW-glutenin and gliadin variations in Tibetan weedrace, Xinjiang rice wheat and Yunnan hulled wheat. *Genet Resour Crop Evol* 49: 327–330.
60. Arbusova V S, Efremova TT, Laikova LI, Maystrenko OI, Popova OM, et al. (1996) The development of precise genetic stocks in two wheat cultivars and their use in genetic analysis. *Euphytica* 89: 11–15.