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

Qian Chen¹ ², Hou-Yang Kang¹ 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 Zhou¹ ² ²*

¹ Triticaceae 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 polyploid 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.

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 AABBBDD [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 *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 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. |
|---------|--------|-----------|--------|---------|----------------|
| Triticum urartu Thum. ex Gandil. | A<sup>u</sup> | TA763 | Lebanon | TUR63A | AF343474 |
| Aegilops bicorns (Forskål) Jaub. et Spach. | S<sup>b</sup> | TA1954 | Egypt | AEB954S | AF343485 |
| Aegilops longissima Schweinf. et Muschl. | S<sup>l</sup> | TA1952 | Israel | AEL9125 | AF343487 |
| Aegilops searsii Feldman et Kislev | S<sup>s</sup> | TA2355 | Israel | AES355S | AF343489 |
| Aegilops sharoneensis Eig | S<sup>sh</sup> | TA2065 | Turkey | AES065S | AF343486 |
| Aegilops speltoides Tausch | S | TA2368 | Turkey | AES368S | AF343483 |
| Aegilops speltoides var. ligustica (Savign.) Fiori | S | TA1770 | Iraq | AEL770S | AF343484 |
| Aegilops tauschii Casson | D | TA60 | Middle East | AET60D | JQ327050 |
| | | TA1691 | Unknown | AET691D | AF343479 |
| Triticum polonicum L. | AB | A5302 | Xinjiang, China | TPO302A | JQ327101 |
| | | A5304 | Xinjiang, China | TPO302B | JQ327102 |
| | | A5304 | Xinjiang, China | TPO304A | JQ327088 |
| | | A5304 | Xinjiang, China | TPO304B | JQ327089 |
| | | PH2209 | Australia | TPO209A | JQ327096 |
| | | | | TPO209B | JQ327097 |
| Triticum turgidum L. | AB | A52233 | Xinjiang, China | TUR233A | JQ327113 |
| | | A52277 | Xinjiang, China | TUR277A | JQ327077 |
| | | A52277 | Xinjiang, China | TUR277B | JQ327078 |
| Triticum durum Desf. | AB | A52349 | Xinjiang, China | TDU349A | JQ327115 |
| | | TDU349B | JQ327116 |
| Triticum durum Desf. cv. Langdon | AB | LDN | USA | TDULA | JQ327057 |
| | | TDULB | JQ327058 |
| Triticum turanicum Jakubz. | AB | A52229 | Xinjiang, China | TTU229A | JQ327109 |
| | | TTU229B | JQ327110 |
| | | TTU279A | JQ327111 |
| | | TTU279B | JQ327112 |
| Triticum dicoccoides (Koern. ex Aschers. et Graeb.) Schweinf. | AB | TAS1 | Israel | TDI51A | AF343481 |
| | | TDI51B | AF343476 |
| | | A5838 | Xinjiang, China | TDI838A | JQ327075 |
| | | TDI838B | JQ327076 |
| Triticum carthlicum Nevski (syn. T. persicum Vav.)AB | PI532494 | Kars, Turkey | TCA494A | JQ327065 |
| | | TCA494B | JQ327066 |
| | | PI532509 | Xinjiang, China | TCA509A | JQ327073 |
| | | TCA509B | JQ327074 |
| Triticum timopheevii (Zhuk.) Zhuk. | AG | TA2 | Armenia | TTI2A | AF343477 |
| | | TTI2G | AF343488 |
| | | PI94761 | Georgia, USA | TTI761A | JQ327126 |
| | | TTI761G | JQ327127 |
| Triticum petropavlovskyi Udacz. et Migusch. | ABD | AS358 | Xinjiang, China | TPE358A | JQ327090 |
| | | TPE358B | JQ327091 |
| | | TPE358D | JQ327092 |
| | | A5359 | Xinjiang, China | TPE359A | JQ327103 |
| | | TPE359B | JQ327104 |
| | | TPE359D | JQ327105 |
| | | A5360 | Xinjiang, China | TPE360A | JQ327106 |
| | | TPE360B | JQ327107 |
| | | TPE360D | JQ327108 |
| Triticum aestivum L. ssp. tibetanum Shao | ABD | AS1026 | Xizang, China | TTB1026A | JQ327123 |
| Species                                | Genome | Accession | Origin               | Abbrev. | GenBank Ac. No. |
|----------------------------------------|--------|-----------|----------------------|---------|-----------------|
| *Triticum aestivum* ssp. *yunnanense* King | ABD    | AS1027    | Xizang, China        | TTB1026B| JQ327124        |
|                                        |        | AS1027D   |                      | TTB1026D| JQ327125        |
|                                        |        | AS1027   | Yunnan, China        | TTB1027A| JQ327062        |
|                                        |        | AS1027B  |                      | TTB1027B| JQ327063        |
|                                        |        | AS1027D  |                      | TTB1027D| JQ327064        |
| *Triticum aestivum* L. ssp. *yunnanense* King | ABD    | AS338     | Yunnan, China        | TYU331A | JQ327131        |
|                                        |        | AS338B    |                      | TYU331B | JQ327132        |
|                                        |        | AS338D    |                      | TYU331D | JQ327133        |
| *Triticum aestivum* L. cv. *Chinese Spring* | ABD    | PI347852  | Switzerland          | TPL852A | JQ327098        |
|                                        |        | PI347852B |                      | TPL852B | JQ327099        |
|                                        |        | PI347852D |                      | TPL852D | JQ327100        |
|                                        |        | PI34785A  | Switzerland          | TPL858A | JQ327120        |
|                                        |        | PI34785B  |                      | TPL858B | JQ327121        |
|                                        |        | PI34785D  |                      | TPL858D | JQ327122        |
| *Triticum compactum* Host              | ABD    | PI124298  | Unknown              | TCO298A | JQ327070        |
|                                        |        | PI124298B |                      | TCO298B | JQ327071        |
|                                        |        | PI124298D |                      | TCO298D | JQ327072        |
|                                        |        | PI352299  | Switzerland          | TCO299A | JQ327067        |
|                                        |        | PI352299B |                      | TCO299B | JQ327068        |
|                                        |        | PI352299D |                      | TCO299D | JQ327069        |
| *Triticum aestivum* L. cv. *Chinese Spring* | ABD    | CS        | Sichuan, China       | TCH5A   | JQ327051        |
|                                        |        | CSB       |                      | TCH5B   | JQ327052        |
|                                        |        | CSD       |                      | TCH5D   | JQ327053        |
| *Triticum aestivum* L. cv. *Chuanrong-16* | ABD    | CN16      | Sichuan, China       | TCN16A  | JQ327054        |
|                                        |        | CN16B     |                      | TCN16B  | JQ327055        |
|                                        |        | CN16D     |                      | TCN16D  | JQ327056        |
| *Triticum aestivum* L. cv. J-11        | ABD    | J-11      | Sichuan, China       | TJ11A   | JQ327079        |
|                                        |        | J11B      |                      | TJ11B   | JQ327080        |
|                                        |        | J11D      |                      | TJ11D   | JQ327081        |
| Synthetic hexaploid wheat             | ABD    | SHW-DPW   |                      | SHWDA   | JQ327059        |
|                                        |        | SHWDB     |                      | SHWDB   | JQ327060        |
|                                        |        | SHWDD     |                      | SHWDD   | JQ327061        |
| *Psathyrostachys juncea* (Fischer) Nevski | Ns      | PI222050  | Afghanistan          | PJU050N | FJ711031        |

The Genebank with AF numbers are from Huang et al. [25], those with JQ numbers have been assinged in this study.
doi:10.1371/journal.pone.0071139.t001
morbidity 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 [30–32] as well as in the identification of parents of allopolyploidy taxa [26,33–36]. Genes such as acetyl-CoA carboxylase 1 (ACO1) [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 Triticea 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 Triticea 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 Pgk1-specific primers, PPF1 (5'-CACCTGGGTGTCCTAAGGGTGTT-3') and PPR1 (5'-ACCACCAGTTGTGGTGCGCTAC-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

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
Figure 1. Maximum-likelihood tree from the exon-intron sequences of the \textit{Pgk-1} gene of \textit{T. petropavlovskyi} 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
Evolutionary History of *Triticum petropavlovskyi*

![Phylogenetic tree showing the evolutionary relationships of *Triticum petropavlovskyi* and other wheat species.](image-url)
Evolutionary History of *Triticum petropavlovskyi*

**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 higher 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 present at position 1220–1308 and had a 89 bp insertion in the G genome of *T. timopheevii*.

**Phylogenetic analyses**

Using *Poastrachys 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 (–Likelikelihood = 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...
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. petropavlovskyi accessions and from T. aestivum cv. Chinese Spring formed a group with 90% bootstrap value and 100% posterior probabilities support. One T. petropavlovskyi 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. petropavlovskyi 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. petropavlovskyi, eleven of T. aestivum and one of Ae. tauschi (TA1691), with 72% bootstrap support. The second one included SHW-DPW and Ae. Tauschi (AS60) with 87% BS and 100% PP.

ML analysis of the intron data yielded a single phylogenetic tree (Lnlikelihood = 1804.87), with the following 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. petropavlovskyi were separated from other Triticum. In the A genome subclade, T. petropavlovskyi (AS359) mapped together with T. polonicum (AS304) (78% BS and 100% PP).

Exon data generated a single ML phylogenetic tree (Lnlikelihood = 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 (PI352509) and T. petropavlovskyi (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. petropavlovskyi grouped together (64% BS and 93% PP), separated from other accessions. Ae. tauschi (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. petropavlovskyi, 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. petropavlovskyi were included in subgroups I and II. T. petropavlovskyi (AS358) and T. polonicum (AS302) were placed at a central branching point. Meanwhile, in the B-type cluster, three accessions of T. petropavlovskyi grouped together in subgroup III, and T. polonicum (AS304) together with SHW-DPW in subgroup IV. In the D-type cluster, T. petropavlovskyi (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 amphidiploid SHW-DPW and Ae. tauschi 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 haplotype 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. petropavlovskyi (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. mhac (PI278660) appeared related to haplotypes of the B genome accessions. Secondly, Ae. tauschi (TA1691) showed a close haplotype relationship with the S genome of Ae. speltoides var. lugistica.

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. petropavlovskyi 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. petropavlovskyi and its putative diploid genome donor, T. polonicum and Ae. tauschi, 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. petropavlovskyi and T. polonicum was 0.68–0.91 MYA for A genome. The divergence time of T. petropavlovskyi and T. polonicum and the B genome was 0.34–0.78 MYA. The divergence time of T. petropavlovskyi 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).
the tree, the tetraploid wheat *T. polonicum* diverged earlier than *T. petropavlovskyi*. In the A genome, the divergence time of *T. petropavlovskyi* (AS358) was later than the other two accessions of *T. petropavlovskyi*. 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. petropavlovskyi* was later than other three Chinese endemic landraces. Between *T. petropavlovskyi* 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. petropavlovskyi*.

**Discussion**

Relationships between *T. petropavlovskyi* 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. petropavlovskyi* 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. petropavlovskyi* 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. petropavlovskyi* 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. petropavlovskyi* are basically shared with *T. spelta*, *T. compactum* and *T. sphaerococcum*. When the B genome is considered, *T. petropavlovskyi* groups in one subclade, comparatively distantly related to *T. aestivum*. In addition, the B-type of MJ network shows that the accessions of *T. petropavlovskyi* (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. petropavlovskyi* in B and D genomes.

Relationships between *T. petropavlovskyi* and the tetraploid wheats

Morphologically, the spikelet of *T. petropavlovskyi* are similar to those of *T. turgidum* and *T. polonicum* [7, 26]. Moreover, the cytology of interspecific hybrids between *T. petropavlovskyi* 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. petropavlovskyi* is more closely related to *T. polonicum* than to *T. durum* or *T. turgidum*. However, Arbusova et al. [60] and Efremova et al. [18] report that the genes supporting the elongated glumes in *T. polonicum* and *T. petropavlovskyi* 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. petropavlovskyi*, while the TCS analysis of exon data indicates that the haplotype of A genome in *T. petropavlovskyi* (AS360) is more closely related to *T. carthilicum* than to other wheats. The phylogenetic analyses and MJ network specific for the B genome showed that three accessions of *T. petropavlovskyi* group together, and are topologically distant from those of tetraploid wheats. This finding indicates that the B genome of *T. petropavlovskyi* diverge from the one of tetraploid species. Yen et al. [3] and Chen et al. [23] also recognized cytologically that the B genome of *T. petropavlovskyi* was different from those of hexaploid wheats.

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

Based on RFLPs, Ward et al. [5] found that *T. petropavlovskyi* is genetically more closely related to accessions of *Ae. tauschii* from Iran than from China. Yang et al. [10] concluded that *T. petropavlovskyi* is derived from a hybrid between *Ae. tauschii* and a presumed *T. polonicum* genotype. However, the phylogenetic analysis of the *Ace-1* genes indicates that the D genome of *T. petropavlovskyi* 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 group with SHW-DPW, *T. compactum* and *T. aestivum* cv. Chiinong-16, while *T. petropavlovskyi*, *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. petropavlovskyi* 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 support that the D genome of *T. petropavlovskyi* is similar to D genome orthologs of *T. aestivum* and only distantly related to *Ae. tauschii*.

The divergence time of the *T. petropavlovskyi*

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. petropavlovskyi* from those of other wheats is dated in this paper is from 0.16 (95% C.I., 0.0–0.38) 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. petropavlovskyi*, 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. petropavlovskyi* 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. petropavlovskyi* and hexaploid species. The divergence time results indicate that *T. polonicum* may have played a role in the evolutionary history of *T. petropavlovskyi*.

The possible origin of *T. petropavlovskyi*

This study shows that the *Pgk-1* sequences of the A genome of *T. petropavlovskyi* 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
amphiploid SHW-DPW. The Pgk-1 B genome data indicate that *T. petropavlovskyi* 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. petropavlovskyi* and *T. polonicum* have very similar A and B genomes. We report a distant relationship between *T. petropavlovskyi* and *Ae. tauschii*. We conclude that *T. petropavlovskyi* is neither derived from an independent allopolyploidization event nor from a single mutation in *A. aestivum*. It is most likely that *T. petropavlovskyi* has an origin starting with a natural cross between *T. aestivum* and *T. polonicum*, with that event taking place around 0.78 MYA.

**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.

**References**

1. Yen C, Yang JL, Luo MC (1985) The origin of the Tibetan weatrace of homoeocline wheat, Chinese Spring, Chengdu guangtou 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.
2. Chen Q, Sun YZ, Dong YS (1985) Cytogenetical studies on interspecific hybrids of Xinjiang wheat. Acta Agron Sin 11: 23–28.
3. Yao JX (1983) Research on a new species in *Triticum*-Xinjiang wheat with rice-like spike. Hereditas (Beijing) 5: 17–20.
4. Kim HS, Ward RW (2000) Pattern of AFLP-based genetic diversity in germplasm pools of common wheat with different geographical or breeding program origins. Euphytica 115: 197–208.
5. Ward RW, Yang ZL, Kim HS, Yen C (1998) Comparative analyses of AFLP diversity in landraces of *Triticum aestivum* and collections of *Aegilops tauschii* from China and South Asia. Theor Appl Genet 96: 312–318.
6. Jakubtsiner MM (1959) K poznaniyu pshenits Kitaja/A contribution to the cytological studies of dwarfing polish wheat (*Triticum aestivum* L.) and *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. Proc Natl Acad Sci 99: 8133–8138.
7. Huang SX, Srikirachornkit A, Forst JD, Xu JX, 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: 803–820.
8. Huang SX, Srikirachornkit A, Xu JX, Fairs 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.
9. Golovinka KA, Glushkov SK, Blinov AG, Mayorov VI, Adikson LR, et al. (2007) Molecular phylogeny of the genus *Triticum*. Plant Syst Evol 264: 195–216.
10. Yang WY, Yen C, Yang JL (1992) Cytogenetic study on the origin of some Chinese wheat landraces. Euphytica 123: 287–293.
11. Dorofeev VF, Filatenko AA, Migushova EF, Udaczin RA, Jakubziner MM (1970) Novoe v poznanii roda *Triticum*-Novace obshchestvo shvestvov. Venstnik S Kh.Nauki 9: 20–24.
12. Riley R, Coucal H, Chapman V (1967) Chromosomal interchanges and the phylogeny of wheat. Heredity 22: 233–247.
13. Shao QQ, Li CS, Basang CR (1980) Semi-wild wheat from Xinzang (Tibet). Acta Genet Sin 7: 150–156.
14. Yang WT, Yen C, Yang JL (1992) Cytogenetic study on the origin of some special Chinese landraces of common wheat. Wheat Inform Serv 75: 14–20.
15. Dorofeev VF, Filatenko AA, Migushova EF, Udaczin RA, Jakubziner MM (1970) Flora of Cultivated plants. In Wheat Dorofeev VF, Migushova EKolo, eds, pp 1–30.
16. Watanabe N, Imamura I (2002) The inheritance and chromosomal location of a gene for long glume phenotype in *Triticum petropavlovskyi* Udzac. et Migusch. J Genet Breed 57: 221–227.
17. Watanabe N, Sekiya T, Sugiyama K, Yamaqishi Y, Imamura I (2002) Telosomic mapping of the homoeologous genes for the long glume phenotype in tetraploid wheat. Euphytica 128: 129–134.
18. Wang HY, Huang XQ, Roder MS, Borner A (2002) Genetic mapping of loci determining long glumes in the genus *Triticum*. Euphytica 123: 267–293.
19. Yen C, Yang JL, Liu XD, Li LR (1983) The distribution of *Aegilops tauschii* Coossen in China with reference of the origin of the Chinese common wheat. In Proceedings of the 7th International Wheat Genetics Symposium, Sakamoto S, ed. Kyoto, Japan, pp 55–58.
20. Yang WT, Zhou YH, Zheng YL (2001) Analysis on chromosome G-band of dwarf polishe wheat (*Triticum polonicum*). J Sichuan Agric Univ 11: 192–114.
21. Liu GX, Zhou YH, Zheng YL, Yang RW, Ding CB (2002) Morphological and cytological studies of dwarfling polishe wheat (*Triticum turgidum* ssp. *politicum*.) from Xinjiang, China. J Sichuan Agric Univ 20: 189–193.
22. Efremova TT, Mastrenko OL, Lyakoova LI, Arbuzova VS, Popova OM (2000) Comparative genetic analysis of hexaploid wheats *Triticum petropavlovskyi* Udzac. et Migusch. and *Triticum aestivum* L. Russ J Genet 36: 1142–1148.
23. Chang PD, Liu DJ, Pri GZ, QJ LL, Huang I, (1980) 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 175–179.
24. Akond ASMGM, Watanabe N (2005) Genetic variation among Portuguese landraces of "Arrancada" wheat and *Triticum petropavlovskyi* by AFLP-based analysis. Genet Resour Crop Evol 52: 619–628.
25. Kang HY, Yang Y, Yuan HJ, Jiang Y, Zhou YH (2008) A new synthesized hexaploids, derived from dwarfin polishe wheat (*Triticum polonicum* L.) and *Aegilops tauschii* Coossen. Intern J Agric Res 3: 252–260.
26. Kang HY, Fan X, Zhang HQ, Sha LN, Sun GL, et al. (2010) The origin of *Triticum petropavlovskyi* Udzac. et Migusch. of the utility of the genes encoding plastid acetyl-CoA carboxylase sequence. Mol Breed 25: 301–305.
27. Sang T (2002) Utility of low-copy nuclear gene sequences in plant phylogenetics. Crit Rev Biochem Mol Biol 37: 121–147.
28. Smith J, Finne M, Woo V (2006) A duplication of gyc predeates divergence within tribe Coronanthereae (Gesneriaceae): phylogenetic analysis and evolution. Plant Syst Evol 261: 243–256.
29. Chahboka D, Lee HY, Faris JD, Croxard A, Challinor B, et al. (2008) Are homoeocloei and the evolution of wheat genomes. Proc Nail Acad Sci 105: 9691–9696.
30. Jiang HY, Wang Y, Yuan HJ, Jiang Y, Zhou YH (2008) A new synthesized holoploid wheat (*Triticum aestivum* L.) and *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. Proc Nail Acad Sci 99: 8133–8138.
31. Chalupska D, Lee HY, Faris JD, Evrard A, Chalhoub B, et al. (2008) Acc homoeoloci and the evolution of wheat genomes. Plant Syst Evol 261: 243–256.
32. Chen Q, Sun YZ, Dong YS (1985) Cytogenetical studies on interspecific hybrids of Xinjiang wheat. Acta Agron Sin 11: 23–28.
33. Yang WT, Zhou YH, Zheng YL, Hua C (2000) Genetic differences and the relationship of gladium between Triticum polonicum and *Triticum petropavlovskyi*. J Triticaceae Crops 20: 1–5.
34. Chen QP (1999) Discussion on origin of Chinese endemic wheat. Guizhou Agri Sci 27: 29–33.
35. Grönroos NP (2005) Comparative genetic analysis of a base for wheat taxonomy revision. Czech J Genet Plant Breed 41: 52–55.
36. Akond ASMGM, Watanabe N, Furuta Y (2008) Comparative genetic diversity of *Triticum aestivum*-*Triticum petropavlovskyi* introgression lines with long and short *Triticum petropavlovskyi* by AFLP-based assessment. Genet Resour Crop Evol 55: 133–142.
41. Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12: 14–15.

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 (Triticaceae; 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) MEGAS: 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 (2003) MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 19: 159–169.

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, Saeidi H, Blatner FR (2011) Phylogeny and genetic diversity of D-genome species of Aegilops and Triticum (Triticaceae, 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. Huaon 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 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, Geis 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 wheedrace, Xinjiang rice wheat and Yunnan hulled wheat. Genet Resour Crop Evol 49: 327–330.

60. Arbuzova V S, Efremova TT, Laikova LJ, Maystruko OL, Popoca OM, et al. (1996) The development of precise genetic stocks in two wheat cultivars and their use in genetic analysis. Euphytica 89: 11–15.