

# Phylogeny of a Genomically Diverse Group of *Elymus* (Poaceae) Allopolyploids Reveals Multiple Levels of Reticulation

Roberta J. Mason-Gamer\*

Department of Biological Sciences, The University of Illinois at Chicago, Chicago, Illinois, United States of America

## Abstract

The grass tribe Triticeae (=Hordeae) comprises only about 300 species, but it is well known for the economically important crop plants wheat, barley, and rye. The group is also recognized as a fascinating example of evolutionary complexity, with a history shaped by numerous events of auto- and allopolyploidy and apparent introgression involving diploids and polyploids. The genus *Elymus* comprises a heterogeneous collection of allopolyploid genome combinations, all of which include at least one set of homoeologs, designated **St**, derived from *Pseudoroegneria*. The current analysis includes a geographically and genomically diverse collection of 21 tetraploid *Elymus* species, and a single hexaploid species. Diploid and polyploid relationships were estimated using four molecular data sets, including one that combines two regions of the chloroplast genome, and three from unlinked nuclear genes: phosphoenolpyruvate carboxylase,  $\beta$ -amylase, and granule-bound starch synthase I. Four gene trees were generated using maximum likelihood, and the phylogenetic placement of the polyploid sequences reveals extensive reticulation beyond allopolyploidy alone. The trees were interpreted with reference to numerous phenomena known to complicate allopolyploid phylogenies, and introgression was identified as a major factor in their history. The work illustrates the interpretation of complicated phylogenetic results through the sequential consideration of numerous possible explanations, and the results highlight the value of careful inspection of multiple independent molecular phylogenetic estimates, with particular focus on the differences among them.

**Citation:** Mason-Gamer RJ (2013) Phylogeny of a Genomically Diverse Group of *Elymus* (Poaceae) Allopolyploids Reveals Multiple Levels of Reticulation. PLoS ONE 8(11): e78449. doi:10.1371/journal.pone.0078449

**Editor:** Jonathan H. Badger, J. Craig Venter Institute, United States of America

**Received:** June 25, 2013; **Accepted:** September 11, 2013; **Published:** November 1, 2013

**Copyright:** © 2013 Roberta J. Mason-Gamer. 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:** The work was supported by National Science Foundation DEB-9974181 (<http://www.nsf.gov>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The author has declared that no competing interests exist.

\* E-mail: [robie@uic.edu](mailto:robie@uic.edu)

## Introduction

In the simplest scenario, an allopolyploid genome represents an additive combination of genomes from its contributing diploid species, so the genomic complement of an allopolyploid individual should be identifiable on a gene tree that includes representatives of all of the potential donor species. In reality, however, there are many factors that complicate the phylogenetic characterization of allopolyploids. Some are inherent in the choice of molecular marker. The chloroplast genome typically represents only the maternal parent in angiosperms, and can't be used to identify the full genomic complement of a polyploid. The highly repetitive internal transcribed spacers (ITS) of the nuclear ribosomal array have the potential to reveal multiple genome donors [1-4], but they are subject to concerted evolution in polyploid genomes [5], so that one or more of the progenitors' arrays might be unrepresented on an ITS gene tree.

These problems can be largely avoided with the use of low-copy nuclear genes; they are biparentally inherited and are less likely to undergo gene conversion, so their gene trees have the potential to reveal all of an allopolyploid's genome donors. As a result, phylogenetic analyses of low copy nuclear genes have been fruitful for clarifying the origins of a wide assortment of allopolyploid species [6-24]. However, there are numerous phenomena that can confound the identification of allopolyploid individuals in phylogenetic analyses of low-copy genes. For example, polyploidization can be associated with extensive genome reorganization, including homoeolog loss (e.g., [18,25-29]); as a result, one or more homoeologs can be missing from a given gene tree, and an allopolyploid's origins can be obscured. The effect will vary across gene trees depending on whether the maternal, paternal, or neither copy is lost. Single-gene duplication, leading to two or more paralogous copies, can also cause misleading phylogenetic estimates. Paralogy is potentially detectable, and is even quite

informative in its own right if gene copies are thoroughly sampled [20,30-37]. However, unsuspected paralogy coupled with limited within-individual sampling and/or paralog loss can result in a confusing gene tree on which each allopolyploid individual is represented by a random selection of one or more paralogs. The identification of a polyploid's progenitors can be further confounded if the relevant diploids are missing from the analysis. In these cases, a polyploid will exhibit gene copies with no clear origin among the diploids. The causes for missing diploids include sparse sampling, and the more intractable problem where one or more of a polyploid's donors are undiscovered or extinct (e.g., [8,20,23,38-41]).

Speciation following polyploidization will yield monophyletic clades of polyploids nested within the donor groups, but subsequent hybridization among them [22,42] will lead to incongruence within those clades across trees. Polyploid origins can be even more difficult to unravel when allopolyploid species form recurrently (e.g., [10,18,43-45]). Subsequent hybridization among them [19,20,22,46-50] can yield individuals with alleles indirectly acquired from many different diploid individuals. Introgression among diploid species prior to polyploidization [40,47,49,51,52] can lead to similar outcomes, in which an allopolyploid has obtained gene copies representing multiple diploid donors. Finally, although a difference in ploidy level is often viewed as an effective barrier to gene exchange, there have been numerous reports of introgression from diploids to polyploids [48,49,53-56], which can further enrich a polyploid's genome while obscuring its phylogenetic history. All of these scenarios - introgression among polyploids, among diploids prior to polyploidization, and among ploidy levels - can lead to conflicting placement of a polyploid's homoeologs on one or more gene trees; thus, conflict can be difficult to interpret in terms of specific phenomena [8,57].

This study focuses on *Elymus*, a diverse, polyphyletic assemblage of allopolyploid species in the wheat tribe, Triticeae Dumort. (A recently-discovered use of the name Hordeae in 1820 - I. I. Martinov, *Tekhnobot. Slovar.*: 314, Aug 1820 - has priority over Triticeae - B. C. J. Dumortier, *Observ. Gramin. Belg.*: 82, Jul-Sep 1824. I use the prevailing name Triticeae here, pending the outcome of a planned proposal to retain its use [M. E. Barkworth, pers. comm.]) While wheat is the most familiar allopolyploid in the tribe (and perhaps among plants), polyploidy has impacted the history of the entire group - about 75% of the approximately 300 species are of polyploid origin [58]. Cytogenetic analyses reveal that the genus *Pseudoroegneria* (genome designation **St**) is a prolific contributor to the tribe's allopolyploids. The **St** genome is found in combination with genomes from numerous other genera, including *Hordeum* (genome **H**), *Australopyrum* (**W**), *Agropyron* (**P**), and an unknown donor with genome designation **Y**. This heterogeneous group of allopolyploid combinations is often collectively classified as *Elymus* sensu lato, as reflected in numerous major floristic works [59-62]. Under this definition, *Elymus* comprises **StStHH** and **StStYY** tetraploids, and **StStStStHH**, **StStHHHH**, **StStStStYY**, **StStYYYY**, **StStHHYY**, **StStYYWW**, and **StStYYPP** hexaploids [58,63-78]. This group, while united by their shared **St** genome,

is obviously polyphyletic, and there has been recent work toward updating their classification to better reflect differences in genomic content [79-82]. There are some additional, higher-order polyploids that include the **St** genome but are not classified as *Elymus* s.l. For example, *Pascopyrum smithii* (Rydb.) Á.Löve combines the **St** and **H** genomes with the **Ns** genome of *Psathyrostachys* in an **StStHHNsNsNs** configuration [83]. The **St** genome also co-occurs with the **E** and/or **J** genomes of *Thinopyrum* [84-86] in complex hexa- and decaploid configurations that might also include full or partial *Taeniatherum*, *Crithopsis*, *Aegilops*, and/or *Dasyphyrum* genomes [87,88].

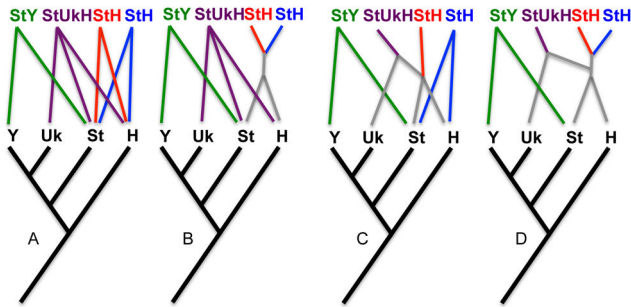
The present work is focused primarily on tetraploid species along with one allohexaploid. As a starting framework, the included species are provisionally divided into four groups (Figure 1). The first two groups include the Eurasian and North American **StStHH** tetraploids, respectively. Their allopolyploid origins from *Hordeum* and *Pseudoroegneria* were initially revealed through cytogenetic analyses (e.g., [66,89-92]) and confirmed in molecular phylogenetic analyses [93-97]. The third group includes **StStYY** tetraploids (= *Roegneria* as used by Yen et al. [81]) which comprise 30-40 species distributed throughout central Asia [73]. Cytogenetic [66,68,71-74,76-78,98-101] and molecular phylogenetic analyses [40,102,103] clearly show the allotetraploid nature of these species, and reveal *Pseudoroegneria* as a genome donor (**St**), but they have been largely inconclusive about the identity of the **Y**-genome donor. The fourth group includes six individuals of *Elymus repens* L. Gould, the sole hexaploid species ( $2n=42$ ) included in the study. The species is native to Europe and Asia, but has become a widespread weed throughout much of the United States and eastern Canada [104]. Cytogenetic analyses of its genome content based on chromosome pairing [105-110] and genomic in situ hybridization [111] are equivocal individually, but taken together, they suggest that *E. repens* has two similar genome sets derived from *Pseudoroegneria* and a third set from *Hordeum* (**StStStStHH**). Molecular phylogenetic analyses provide a more complex picture, suggesting three distinct donors, and introgression from at least one additional genus within the tribe [52,112,113], along with apparent contributions from tribes Bromaceae and Paniceae [52], and Poaceae [113].

The four provisional polyploid groups have been examined individually, and here the relationships among them are investigated using data from two chloroplast DNA regions, and three single-copy nuclear genes. The specific goals are to determine: (1) whether the four provisional groups represent phylogenetically distinct lineages; and (2) whether specific evolutionary phenomena (homoeolog loss, paralogy, progenitor extinction, multiple polyploid origins, and/or introgression) can be identified, or excluded, as explanations for reticulate patterns beyond allopolyploidy itself.

## Materials and Methods

### Data sets

The samples include allopolyploid representatives of *Elymus* from the four provisional groups, along with a broad sample of



**Figure 1. Evolutionary scenarios.** Four scenarios representing separate (a) or partially sequential (b–d) origins of the four provisional categories of polyploids: **StStHH** tetraploids native to North American (blue) and Eurasia (red), Asian **StStYY** tetraploids (green), and *E. repens* hexaploids (purple). Additional scenarios are conceivable, including recurrent origins of each group.

doi: 10.1371/journal.pone.0078449.g001

monogenomic, mostly diploid representatives from throughout the tribe (Table S1). The chloroplast DNA (cpDNA) data set (Dataset S1) includes the trnT-trnL-trnF region, and the rpoA gene. Sequences are derived from several previous analyses (see Table S1 footnotes) with the addition of new sequences representing 16 Eurasian **StStHH** and **StStYY** tetraploid *Elymus* accessions and eight diploid *Pseudoroegneria* accessions (Table S1). The trnT-trnL-trnF sequences were amplified using *a* and *f* primers, and sequenced with *a*, *b*, *c*, *d*, *e*, and *f* primers [114]. The rpoA gene was amplified using rpoA1 and rpoA2 primers, and sequenced with rpoA1, rpoA2, rpoA4, rpoA5, rpoA8, and rpoA9 primers [115]. Amplification and sequencing followed methods previous applied to this tribe [116].

Each of the nuclear gene data sets – phosphoenolpyruvate carboxylase (pepC; Dataset S2),  $\beta$ -amylase (Dataset S3), and granule-bound starch synthase I (GBSSI; Dataset S4) – combine data from numerous prior analyses (see Table S1 footnotes) that were primarily designed to reveal the relationships of the individual polyploid groups to their diploid progenitors. The present analyses were undertaken to further clarify the relationships among these diverse groups. In addition, a phylogenetically broader analysis of GBSSI exon data includes additional representatives of the Pooideae, and a few representatives of the Bambusoideae (Table S2; Dataset S5); this analysis was used to clarify the placement of a divergent clade of sequences from the *E. repens* hexaploids.

### Phylogenetic analyses

Three of the data sets were partitioned for model specification (Table S3). The chloroplast DNA data were partitioned by region: the trnT-trnL-trnF region, which consists primarily of non-coding sequences of the intergenic spacers and the trnL intron, and the rpoA gene, which consists mainly of protein-coding sequence. The  $\beta$ -amylase and starch synthase gene data sets were both partitioned into exons vs. introns. The phosphoenolpyruvate carboxylase sequences

were analyzed under a single model, because over 90% of the data set represents a single intron, with only small portions of exons 1 and 2 (59 and 23 bp, respectively). The GBSSI exon data used in the broader phylogenetic analysis were also analyzed under a single model.

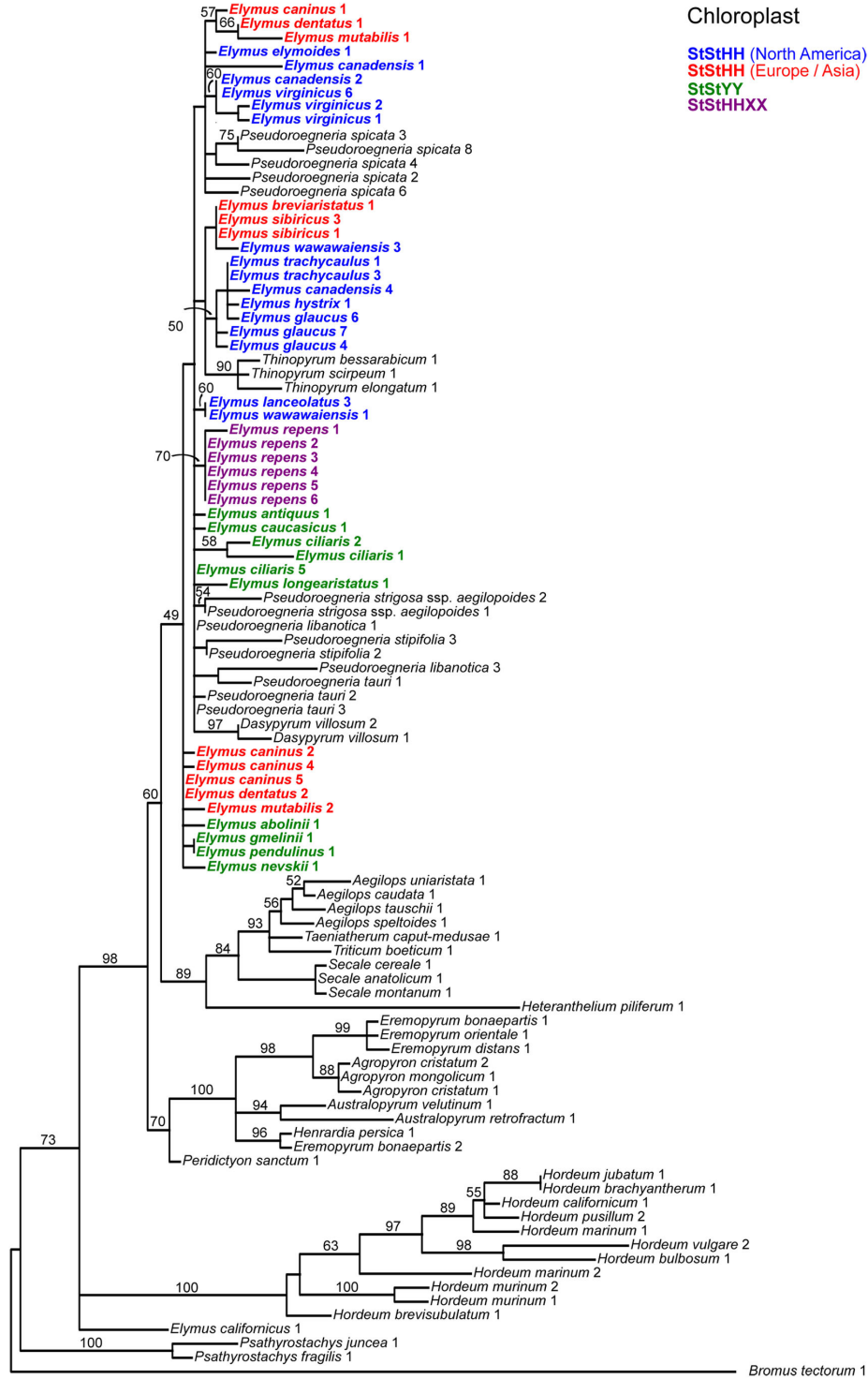
For each partition, jModelTest [117] was used to select among 24 candidate models. These include three substitution models (AC=AG=AT=CG=CT=GT; AC=AT=CG=GT  $\neq$  AG=CT; and AC $\neq$ AG $\neq$ AT $\neq$ CG $\neq$ CT $\neq$ GT), first under an assumption of equal nucleotide frequencies: JC [118], K80 [119], and SYM [120] – and then allowing for unequal nucleotide frequencies: F81 [121], HKY [122], and GTR [123]. Each of these was combined with four among-site rate parameters: equal rates, some proportion of sites invariant (I)[122], gamma-distributed rate variation ( $\Gamma$ )[124], or I+ $\Gamma$  [125,126]. Input files were formatted for jModelTest using ALTER [127]; model parameters were estimated on maximum-likelihood (ML) estimated trees and models were selected using the Akaike information criterion [128].

Maximum-likelihood phylogenetic analyses were run under the selected models using GARLI [129] v. 2.0 (<http://garli.googlecode.com>). Fifty analyses were run for each data set with random starting topologies, for an unlimited number of generations, and automatic termination following 10,000 generations without a meaningful change in score (ln L increase of 0.01). The tree with the best score was chosen to depict each gene tree (Figures 2–6), and for final model parameter estimates (Table S3). Bootstrap support estimates were based on 100 ML replicates under the same model and under the same conditions, except that the change requirement for search termination was increased from 0.01 to 0.02.

## Results

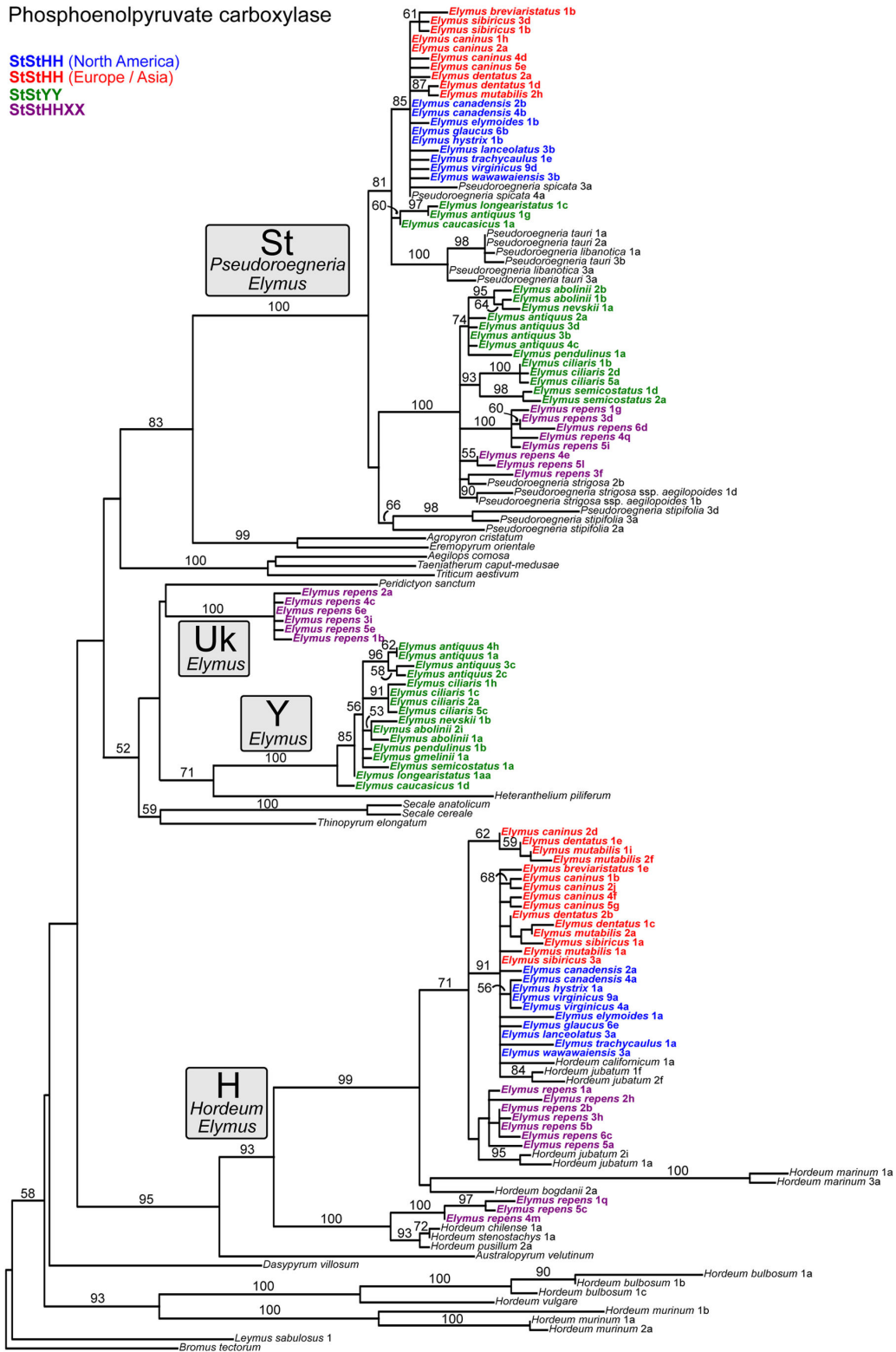
### Chloroplast DNA tree

The cpDNA data (Figure 2) place the sequences from all of the polyploid *Elymus* individuals into a single weakly-supported clade, along with *Dasypyrum*, *Pseudoroegneria*, and *Thinopyrum*. Thus, all species of this fairly diverse, polyphyletic sample of *Elymus* species appear to share the same maternal donor. The consistent placement of *Elymus* with *Pseudoroegneria* on other gene trees (below), and never with *Dasypyrum* or *Thinopyrum*, points to *Pseudoroegneria* as the maternal donor to *Elymus*. *Elymus* terminal branch lengths are short and there is very little resolution among the individuals. The only well-supported groups within the large clade are those that group the two *Dasypyrum villosum* individuals (97%) and the three *Thinopyrum* individuals (90%); the six *E. repens* sequences are grouped with moderate (70%) support. None of the three pre-defined tetraploid groups (Eurasian **StStHH**, North American **StStHH**, or Asian **StStYY**) form clades. The separation of the sole North American *Pseudoroegneria* species (*P. spicata*) from the Eurasian species on the ML tree lacks meaningful support. Thus, although the addition of 24 species of *Elymus* and *Pseudoroegneria* provides a much broader geographic sample than previous analyses, the tree reveals little about the role of geography, or the independent



**Figure 2. Chloroplast DNA gene tree.** The data set (Dataset S1) was partitioned by genome region for *a priori* model specification in jModelTest, and simultaneous parameter estimation in GARLI: the trnT-trnL-trnF region (GTR+ $\Gamma$ ) and the rpoA gene (GTR+ $\Gamma$ ). This represents the best tree from 50 GARLI search replicates. ML bootstrap results are based 100 GARLI replicates under the same models as used in the tree searches. Monogenomic representatives of the tribe are in black font; *Elymus* representatives are in colored, boldface font. Colors distinguish the four hypothetical *Elymus* species groups as described in the text: North American (blue) and Eurasian (red) StStHH tetraploids, Asian StStYY tetraploids (green), and *E. repens* hexaploids (purple).

doi: 10.1371/journal.pone.0078449.g002

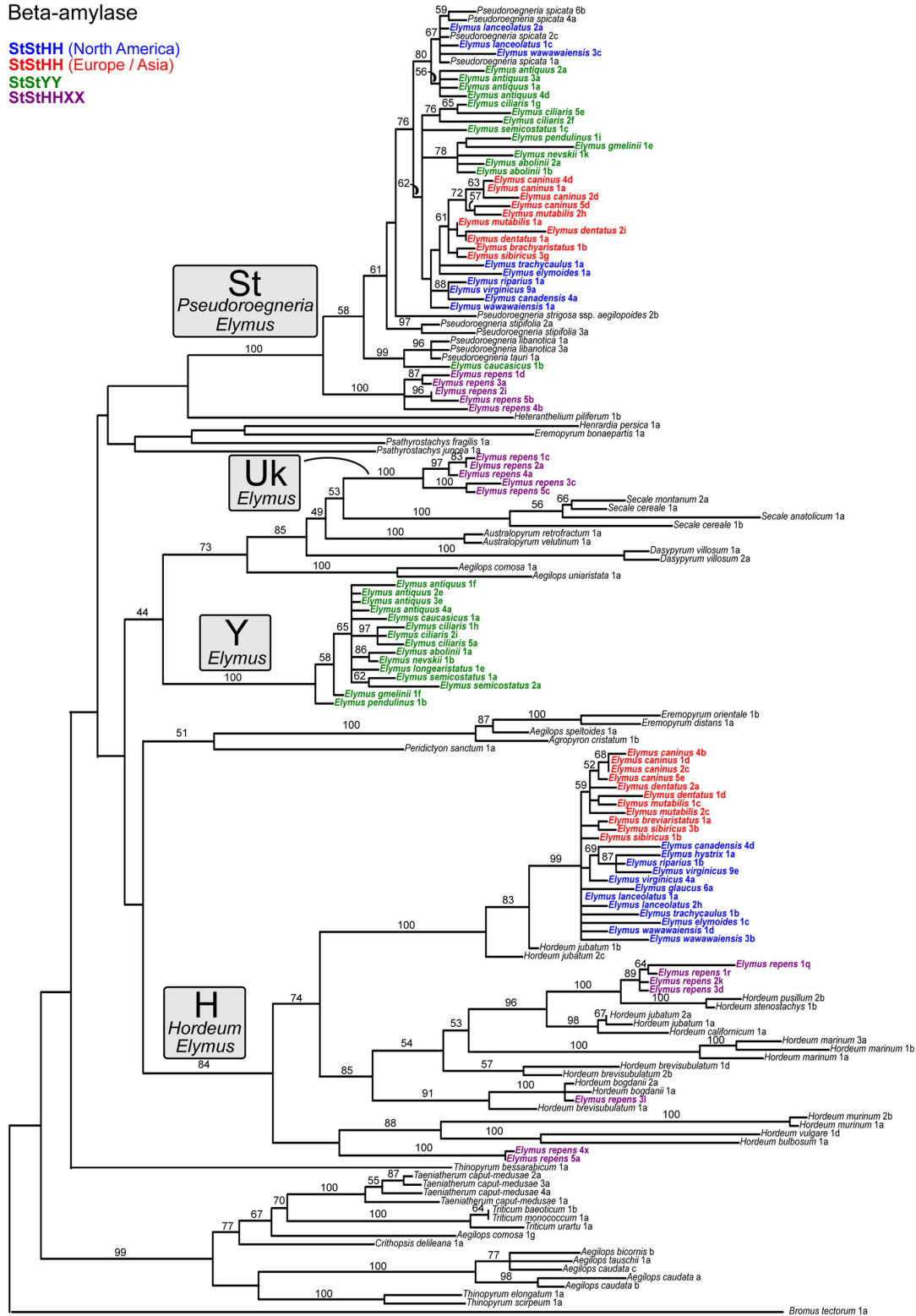


**Figure 3. Phosphoenolpyruvate carboxylase gene tree.** The data set (Dataset S2) was treated as a single partition for *a priori* model specification in jModelTest (HKY+ $\Gamma$ ), and simultaneous parameter estimation in GARLI. This represents the best tree from 50 GARLI search replicates. ML bootstrap results are based 100 GARLI replicates under the same models as used in the tree searches. Font colors follow Figure 2.

doi: 10.1371/journal.pone.0078449.g003

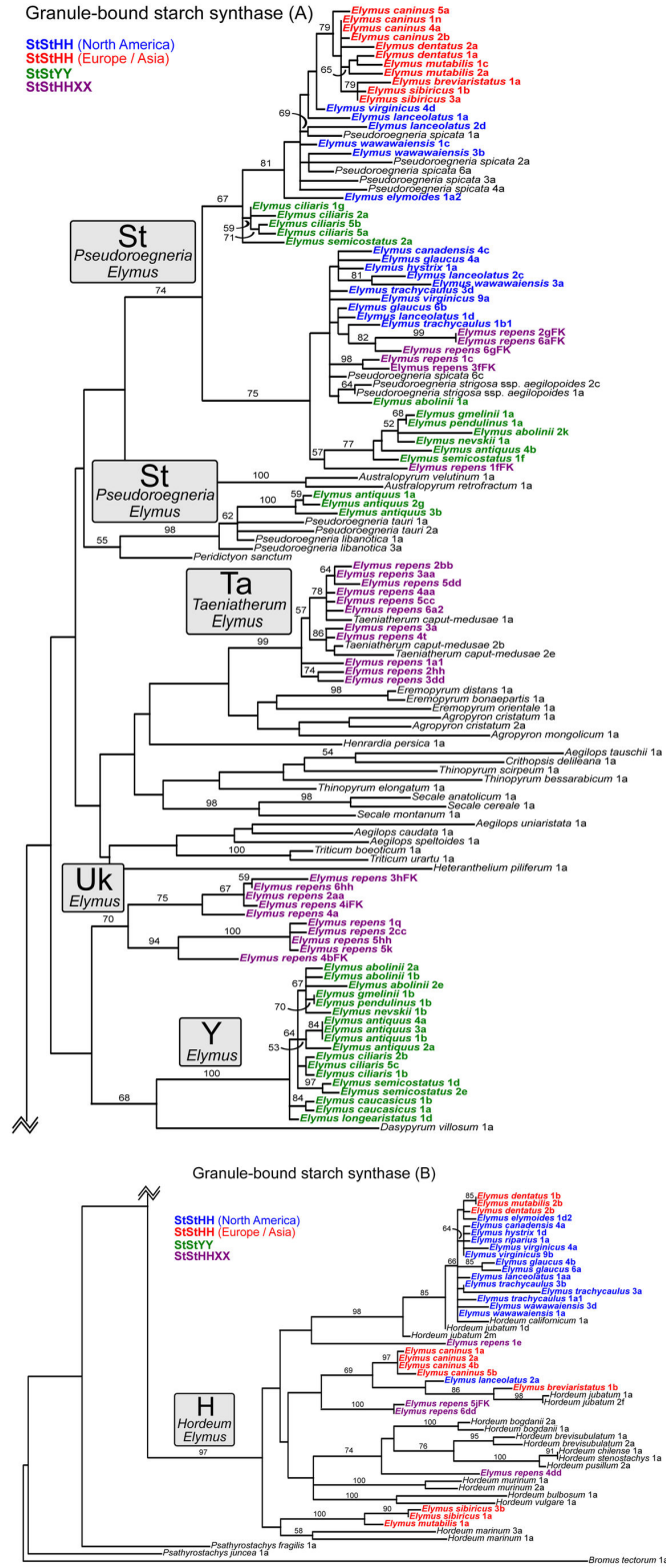
Beta-amylase

StStHH (North America)  
 StStHH (Europe / Asia)  
 StStYY  
 StStHXX



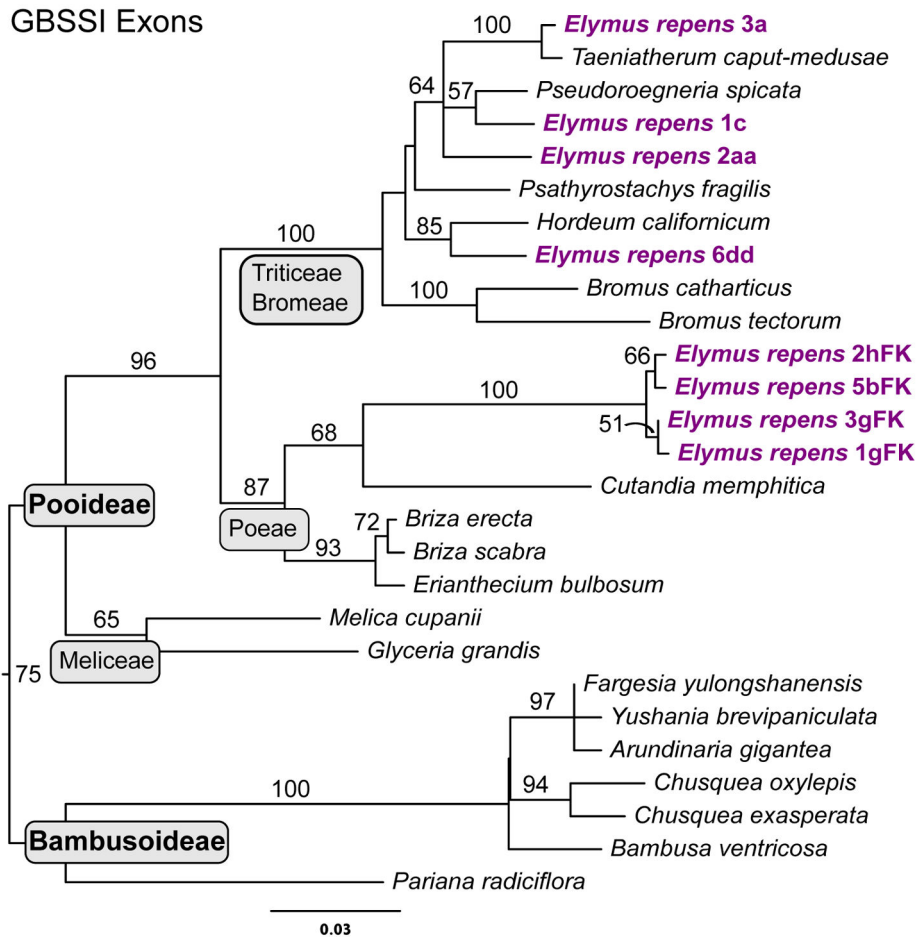
**Figure 4. Beta amylose gene tree.** The data set (Dataset S3) was partitioned for *a priori* model specification in jModelTest, and simultaneous parameter estimation in GARLI: exons (K80+ $\Gamma$ ) and introns (GTR+ $\Gamma$ ). This represents the best tree from 50 GARLI search replicates. ML bootstrap results are based 100 GARLI replicates under the same models as used in the tree searches. Font colors follow Figure 2.

doi: 10.1371/journal.pone.0078449.g004



**Figure 5. Granule-bound starch synthase I gene tree.** The data set (Dataset S4) was partitioned *a priori* model specification in jModelTest, and simultaneous parameter estimation in GARLI: exons (GTR+ $\Gamma$ ) and introns (GTR+ $\Gamma$ ). This represents the best tree from 50 GARLI search replicates. ML bootstrap results are based 100 GARLI replicates under the same models as used in the tree searches. Font colors follow Figure 2.

doi: 10.1371/journal.pone.0078449.g005



**Figure 6. Granule-bound starch synthase I exon tree: Pooids and Bambusoideae.** The data set (Dataset S5) was treated as a single partition for *a priori* model specification in jModelTest (GTR+I+ $\Gamma$ ), and simultaneous parameter estimation in GARLI. ML bootstrap results are based 100 GARLI replicates under the same model. The Triticaceae clade represents a small subset of the individuals in the full GBSSI tree (Figure 5). *Elymus repens* sequences are highlighted in purple, boldface font.

doi: 10.1371/journal.pone.0078449.g006

formation of multiple polyploid combinations, in the patterns of cpDNA differentiation among the species.

### Phosphoenolpyruvate carboxylase

Within the **St** clade on the pepC tree (Figure 3), the sequences from both North American and Eurasian **StStHH** *Elymus* are grouped with *P. spicata* (85%), the only native North American *Pseudoroegneria* species. There is little structure within this clade, and the sequences are very similar to one another. Sequences from *E. repens* and the tetraploid **StStYY** Asian species are grouped with *P. strigosa* (100%), a Eurasian species. The only exception to this involves three **StStYY** species: *E. antiquus* (one out of four individuals), *E. caucasicus*, and *E. longearistatus*; their **St** sequences are closer to (though not within) the “wrong” clade in terms of phylogenetic relationship and genetic similarity. The other three *E. antiquus* individuals are in the main group of **StStYY** species.

The **H**-sequences from the North American and Eurasian **StStHH** species are closely related to one another (as in the **St** clade), and to those from *E. repens*. They group with diploid *H. californicum* and tetraploid *H. jubatum* (71%), which are both North American natives. There is little phylogenetic structure among these sequences, but most of the sequences from the *Elymus* tetraploids are closely related to *H. californicum* and one genome from *H. jubatum*, while *E. repens* is weakly grouped with the other genome of *H. jubatum*. Three sequences, representing *E. repens* individuals 1, 4, and 5, form a separate well-supported clade (100%) outside of the main *Elymus* clade, with *H. chilense* and *H. stenostachys* (both South American), and *H. pusillum* (North American). Two of these individuals, 1 and 5, also have copies in the main clade.

The **Y**-sequences from the **StStYY** individuals form a phylogenetically distinct clade (100%) with a moderately-supported relationship (71%) with *Heteranthelium piliferum*. *Elymus antiquus*, which was polyphyletic in the **St** clade, is strongly monophyletic (96%) in the **Y** clade. An additional



clade, representing a third clade of *E. repens* sequences of unknown origin (Figure 3, “UK”), is phylogenetically distinct from the **St**, **H**, and **Y** sequences. The origin of this clade is unclear; its association with the **Y**-sequence clade and *H. piliferum* is very weak (<50%). The “UK” and **Y** sequences are generally in the same part of the tree, but support for a close relationship between them, and thus any suggestion that the “UK” and **Y** sequences are from the same unknown genome donor, is weak at best.

### Beta-amylase

On the  $\beta$ -amylase tree (Figure 4), the **St** sequences from the **StStHH** and **StStYY** tetraploids form a clade with North American *P. spicata* (76%). The single exception is one *E. caucasicus* sequence with a strong (99%) relationship to three Iranian *Pseudoroegneria* accessions – two of *P. libanotica* and one of *P. tauri*. The **StStHH** *Elymus* / *P. spicata* relationship reflects the results on the pepC tree (Figure 3), but the inclusion of the sequences from the **StStYY** sequences is in sharp contrast. The sequences from Eurasian **StStHH** species form a weak monophyletic subgroup (61%), while those from North American **StStHH** species and from the **StStYY** species are polyphyletic. The *E. repens* **St** sequences form a phylogenetically distinct monophyletic group (100%), sister to the remaining **St** sequences.

The **H** sequences of all of the **StStHH** species form a strong clade (99%) that is closely related (100%) to one of the genomes of the tetraploid North American species *H. jubatum* (sequences 1b and 2c). As on the pepC tree, some of the sequences from *E. repens* form a close relationship (100%) with those from *H. stenostachys* and *H. pusillum*. These in turn group with *H. californicum* and the other *H. jubatum* homoeologs (sequences 1a and 2a; 96%). In sharp contrast to the pepC tree, none of the *E. repens* **H** sequences group with those from **StStHH** tetraploids. One of them (3l) groups (91%) with two accessions of *H. bogdanii* representing Tajikistan and China, and two (4x and 5a) are together (100%) on a long branch that is not closely associated with any of the sampled *Hordeum* species.

The **Y**-genome sequences from the **StStYY** tetraploids again form a phylogenetically distinct clade (100%). In contrast to the pepC tree, this tree shows no evidence of a relationship between the **Y**-genome clade and *H. piliferum*; furthermore, they show no clear association with any of the diploid donors. The *E. repens* “UK” sequences again form a monophyletic group (100%), this time within a larger, moderately-supported (85%) group with *Australopyrum*, *Dasypyrum*, and *Secale*. As on the pepC tree, the “UK” clade shows no apparent close relationship to the **Y**-genome clade of the **StStYY** tetraploids.

### Granule-bound starch synthase

The **St**-genome sequences from *Elymus* show close relationships to three different *Pseudoroegneria* species (Figure 5a). First, the **St** sequences from Eurasian **StStHH** tetraploids, and about half of those from the North American tetraploids, form a monophyletic group (81%) with North American *P. spicata*. Sequences from two of the **StStYY** species, *E. ciliaris* and *E. semicostatus*, are weakly (67%)

linked to this group as well. Second, the remaining sequences from the North American **StStHH** tetraploids, most from the **StStYY** species, and all from hexaploid *E. repens* form a clade with *P. strigosa* and a single *P. spicata* sequence (75%). Within this clade, none of the species groups (**StStHH**, **StStYY**, *E. repens*) are monophyletic. Finally, the **St** sequences from three of the four **StStYY** *E. antiquus* individuals are strongly (98%) grouped with *P. tauri* and *P. libanotica* (four accessions, all representing Iran). These sequences are, at best, only weakly associated with the main **St** group.

The relationships among the **H**-genome sequences from *Elymus* and *Hordeum* (Figure 5b) are complicated. On one hand, the results partly reflect those on the pepC and/or  $\beta$ -amylase trees: most of the sequences from the North American **StStHH** tetraploids, and several from the Eurasian **StStHH** tetraploids, are closely related to the North American *Hordeum* species *H. jubatum* (allotetraploid) and *H. californicum* (diploid). On the other hand, the remaining **H**-genome sequences (including one from a North American tetraploid, most from the Eurasian tetraploids, and all from hexaploid *E. repens*) are scattered among the other *Hordeum* species. Some of the tetraploids' sequences are grouped with the other genome of the allotetraploid *H. jubatum*, suggesting possible involvement between *Elymus* and *Hordeum* at the tetraploid level. The three remaining sequences from the Eurasian tetraploids (*E. sibiricus* 1a and 3b, and *E. mutabilis* 1a) and the *E. repens* sequences are on long branches with no clear relationship to any particular *Hordeum* species.

The **Y**-genome sequences of the **StStYY** tetraploids (Figure 5a) again form a well-supported clade (100%); this tree uniquely suggests a weak relationship (68%) to *Dasypyrum villosum*. The *E. repens* “UK” sequences form a moderately supported group (70%). As on the pepC tree, the *E. repens* “UK” sequences are very weakly associated with the **Y** clade, but again, there is no clear suggestion that they are derived from the same unknown diploid donor.

The GBSSI tree shows a unique, fourth group of *E. repens* sequences very closely associated (99%) with *Taeniatherum* (**Ta**; Figure 5a). All of the *E. repens* individuals are represented in this clade, and no related sequence is found in any of the tetraploids. Thus, the *E. repens* genome harbors four distinct GBSSI copies within the Triticeae. In addition, three *E. repens* individuals have a divergent fifth GBSSI copy type that falls outside of the tribe (Figure 6). The sample of non-Triticeae pooid taxa is limited, but this divergent clade of *E. repens* sequences clearly fall outside of the well-supported Triticeae / Bromeae clade, and apparently within the Poaeae.

## Discussion

### General relationships among lineages

The provisional division of *Elymus* into four groups (North American **StStHH** tetraploids, Eurasian **StStHH** tetraploids, Asian **StStYY** tetraploids, and the Eurasian hexaploid *E. repens*) reflects a simple preliminary hypothesis that the groups represent distinct evolutionary lineages, whether arising independently from one another (Figure 1a) or partially sequentially (Figure 1b–d). The placement of *Elymus*

individuals on each the four gene trees can highlight deviations from the hypothesis that all – or any – of the groups represent straightforward lineages. In summary, the data support three independently-derived lineages (Figure 1b): a combined North American and Eurasian **StStHH** lineage, the **StStYY** species, and *E. repens*. All three, however, show complicated phylogenetic patterns, as discussed below under the heading “Potentially confounding phenomena.”

### The cpDNA tree does not distinguish lineages

Of the four trees, the cpDNA tree (Figure 2) is the least informative with respect to the origins and evolution of the polyploid lineages, except in its suggestion that *E. repens* was derived from one (or multiple similar) maternal donor(s). This is due largely to lack of resolution within the large *Pseudoroegneria* / *Elymus* clade. All three tetraploid groups are polyphyletic on this tree, but with support too weak to convincingly suggest actual polyphyletic origins. This tree, when considered alone, suggests three potential chloroplast donors: *Dasypyrum*, *Pseudoroegneria*, or *Thinopyrum*, but arguments in favor of *Pseudoroegneria* as the maternal have already been presented in detail elsewhere [116]. Briefly (1), there are no cytogenetic or molecular phylogenetic data that point to either *Dasypyrum* or *Thinopyrum* as a contributor to the *Elymus* polyploids; and (2) a *Dasypyrum* / *Thinopyrum* / *Pseudoroegneria* relationship appears on no other phylogenetic estimates to date, suggesting that *Dasypyrum* and *Thinopyrum* are “misplaced” with *Pseudoroegneria* on the cpDNA tree as a result of past introgression.

### The North American and Eurasian **StStHH** tetraploids are united, but with signs of introgression

Some previous studies suggested separate origins of the Eurasian and North American **StStHH** species groups from Eurasian and North American progenitors, respectively [97,130,131]. In contrast, the results here point to a North American origin for both groups. The three nuclear genes, considered individually and in combination, support a close relationship between the Eurasian and North American **StStHH** tetraploid groups, but with evidence of introgression (Figure 7a). The phosphoenolpyruvate carboxylase and  $\beta$ -amylase trees generally unite the **StStHH** tetraploids and are consistent with a North American origin; their **St** genomes are closely related to the North American native diploid *P. spicata*, and their **H** genomes are closely related to the North American species *H. californicum* and tetraploid *H. jubatum* (which itself has a *H. californicum*-like donor [38]). The GBSSI tree is more difficult to interpret, with both the North American *Elymus* **St** sequences, and the Eurasian *Elymus* **H** sequences, polyphyletic. This pattern is seen on only one gene tree, so it is more likely an indicator of GBSSI introgression than of multiple origins.

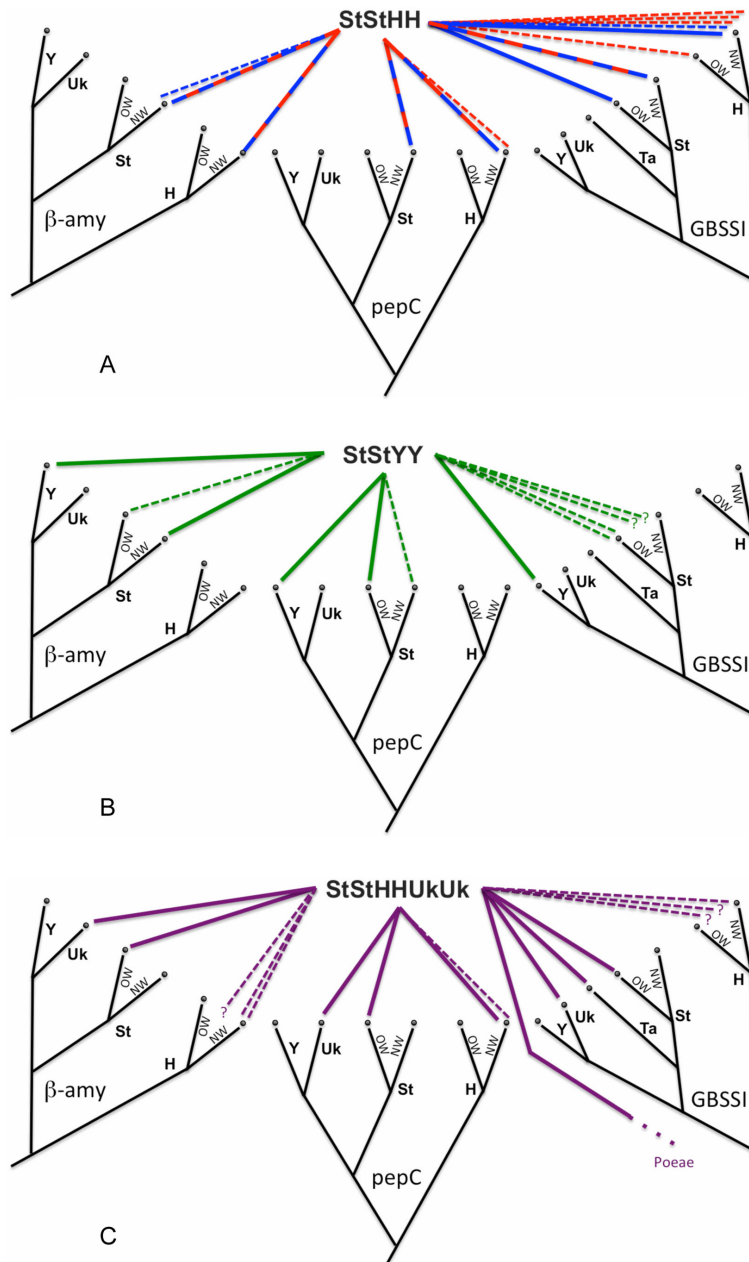
### The Asian **StStYY** tetraploids probably represent a single evolutionary lineage with subsequent introgression

The **Y**-genome of the Asian tetraploids represents a phylogenetically distinct entity. It does not fit into an existing genus, as shown here and in earlier studies [132,133] (although one previously-published gene tree shows a close relationship between **Y**-genome sequences and *Australopyrum* [103]). The hypothesis that the **Y** genome is a derivative of the **St** genome [134,135] is not supported here. Thus, the proposed application of separate generic status (*Roegneria* [81]) to the cytogenetically distinct **StStYY** is further justified by phylogenetic data. The donor of its **St** genome is unequivocally *Pseudoroegneria*, but the nuclear gene trees are in conflict with regard to which species is the most likely donor. While the pepC tree points to the Eurasian species *P. strigosa* as a likely donor, the  $\beta$ -amylase tree instead indicates *P. spicata*, placing the **StStHH** species in a largely North American clade with the **StStHH** species. The sequences are widely polyphyletic on the GBSSI tree, again suggesting a history of introgression (Figure 7b).

*Elymus repens* represents a single, very complicated lineage. In spite of its highly polyphyletic **H**-genome sequences on all three nuclear trees (Figures 3–5, 7c), *E. repens* as sampled here probably represents a single lineage. This is supported by numerous shared genetic characteristics, including a third “unknown” genome (Figures 3–5) and a *Taeniatherum*-like GBSSI allele (Figure 5), features that are also shared by individuals sampled from their native range [52]. Based on its presumed **StStStStHH** genomic content, the hexaploid *E. repens* might conceivably be derived from an existing tetraploid **StStHH** lineage, and with the uncertainties about its donors, it is even difficult to rule out an **StStYY** progenitor a priori. The phylogenetic data are inconsistent. For example, the pepC **H**-genome sequences (Figure 3) clearly link *E. repens* with the **StStHH** tetraploids and North American *Hordeum* species, while the **St** sequences on the same tree group *E. repens* strongly with the **StStYY** species and a Eurasian *Pseudoroegneria* species, *P. strigosa*. The  $\beta$ -amylase **St** and **H** sequence data (Figure 4) separate *E. repens* from both the **StStHH** and the **StStYY** tetraploids, while the GBSSI **St** sequences places it with individuals from both groups. Without a consistent relationship between *E. repens* and the **StStHH** or **StStYY** tetraploids, I tentatively conclude that the origin of *E. repens* is independent of the sampled tetraploids. However, with extensive conflict among data sets, the question is not fully resolved.

### Potentially confounding phenomena

In phylogenetic analyses of allopolyploids and their diploid relatives, non-conflicting relationships across gene trees suggest an uncomplicated evolutionary history, while discordance among trees suggests a history of reticulation beyond allopolyploidization itself, such as introgression among polyploid lineages and/or their progenitors. As outlined in the Introduction, numerous evolutionary phenomena (Table 1) can affect the placement of polyploids on gene trees. These are discussed below with respect to the present results from the



**Figure 7. Summary of *Elymus* genetic diversity.** The black trees represent diploid phylogenies; basal relationships are poorly supported on the actual gene trees, so some relationships among **St**, **Y**, **H**, and “**UK**” are unclear. The colored lines represent contributions to polyploids; unbroken and dotted lines represent major and minor contributions, respectively. Colors follow Figures 1–5. The **St** (*Pseudoroegneria*) and **H** (*Hordeum*) species are distinguished as Old World (OW) and New World (NW). a. **StStHH** species. The red-and-blue lines indicate where the North American (blue) and Eurasian (red) groups received major contributions from the same, or phylogenetically similar, donors. The  $\beta$ -amylase and pepC trees suggest fairly straightforward New World origins, with few minor St- or H-genome contributions. The GBSSI tree shows more complicated origins, especially with respect to *Hordeum*'s contribution. b. **StStYY** species. Without any representative Y-genome diploids, a single donor is hypothesized based on monophyly and sequence similarity (Figures 3–5). Primary St-genome donors are a mix of New World ( $\beta$ -amylase) and Old World (pepC and GBSSI), with minor contributions from the alternate region in all three cases. c. *Elymus repens*. All trees are at least consistent with a single Old World St-genome contribution. The presence of a third (“UK”) clade on all trees reveals an unknown genome donor. There is no single, major H-genome contributor, but a mix of multiple contributors. Contributions unique to the GBSSI tree (from *Taeniatherum*, and from an unknown species from the tribe Poeeae) are consistent with introgression after polyploidization.

doi: 10.1371/journal.pone.0078449.g007

**Table 1.** Phenomena that affect polyploid phylogenies.

Phenomenon	Similar effect on all trees?	Affects diploids or polyploids?	Potential effects	Proposed examples
1. Homoeolog loss following polyploidization	No	Polyploids	Placement of polyploid varies depending on which copy is retained.	<i>Tragopogon</i> [25]; <i>Solanum</i> [18]; <i>Triticum</i> [26]; Review [29]; <i>Brassica</i> [27]; <i>Brassica</i> , <i>Arabidopsis</i> [28].
2. Paralogy; paralogs retained and intensively sampled	No	Both	Intra-individual polymorphism shared by all species with duplication; duplicate clades on tree.	<i>Gossypium</i> [33]; <i>Viola</i> [35]; <i>Paeonia</i> [32,36]; Brassicaceae, Cleomaceae [34]; Rosaceae [30]; Gesneriaceae [31]; <i>Viburnum</i> [37].
3. Paralogy; limited paralog sampling	No	Both	Sporadic intra-individual polymorphism, polyphyletic taxa, incongruence with other gene trees.	–
4. Paralogy; loss of paralogs	No	Both	Sporadic intra-individual polymorphism, polyphyletic taxa, incongruence with other gene trees.	<i>Mentzelia</i> [20]; <i>Glycine</i> [16].
5. Diploid progenitor(s) extinct, undiscovered, or unsampled	Yes	Polyploids	One or more homoeologs unassociated with any diploid.	<i>Hordeum</i> [38]; <i>Mentzelia</i> [20]; <i>Cerastium</i> [8]; <i>Cardamine</i> [11]; <i>Elymus</i> [40]; <i>Microseris</i> [41]; <i>Dryopteris</i> [23].
6. Recurrent origins of a polyploid combination	Yes	Polyploids	Polyploid sequences polyphyletic within donor clades and congruent among gene trees.	<i>Spiranthes</i> [43]; <i>Solanum</i> [18]; <i>Tragopogon</i> [44]; <i>Rosa</i> [10]; Review [45].
7. Introgression among genomically similar recurrent polyploids	No	Polyploids	Polyploid sequences polyphyletic within donor clades and incongruent among gene trees.	<i>Centauria</i> [47]; <i>Dactylorhiza</i> [48]; <i>Viola</i> [22].
8. Introgression among species derived from a single polyploid ancestor	No	Polyploids	Polyploid sequences monophyletic within donor clades and otherwise incongruent among gene trees.	<i>Spartina</i> [42]; <i>Viola</i> [22].
9. Introgression among genomically distinct polyploid lineages	No	Polyploids	Polyploid sequences broadly polyphyletic across multiple donor clades and incongruent among gene trees.	<i>Galeopsis</i> [19]; <i>Mentzelia</i> [20]; <i>Houstonia</i> [46]; <i>Arabidopsis</i> [49]; <i>Elymus</i> [50].
10. Introgression from diploid to polyploid	No	Polyploids	Polyploid sequences polyphyletic within donor clades and incongruent among gene trees.	<i>Senecio</i> [53]; <i>Dactylorhiza</i> [48]; <i>Arabidopsis</i> [49]; <i>Achillea</i> [54]; <i>Dodecatheon</i> [55]; <i>Betula</i> [56].
11. Introgression among diploids prior to polyploidization	No	Both	Sequences from diploids and their derived polyploids are polyphyletic and incongruent among gene trees.	<i>Glycine</i> [51]; <i>Centauria</i> [47]; <i>Arabidopsis</i> [49]; <i>Elymus</i> [52]; <i>Elymus</i> [40].

doi: 10.1371/journal.pone.0078449.t001

wheat tribe analyses, with the caveat that they can be difficult to pin down individually, not only because some have similar outcomes, but also because they can act simultaneously and obscure one another's effects.

Multiple homologous gene copies (Table 1, 1–4) can mislead phylogenetic analyses if they are unsuspected, incompletely sampled, or have undergone copy losses. These include homoeologs (the homologous gene copies representing different genome sets in a polyploid), and paralogs (used here to refer only to copies arising through within-genome gene duplication). Homoeolog loss following polyploidization (Table 1, 1) has occurred rarely, if at all, for the genes examined. Occasional homoeologs are missing from some taxa (NR in Table S1), but at least for these three genes, copy loss is clearly not widespread in *Elymus*. The occasional missing copies might represent individual instances of copy loss, failure to amplify due to changes at a priming site, or sampling artifacts; the last of these is the most straightforward explanation.

Paralogy (Table 1, 2–4) can either complicate or clarify phylogenetic relationships. If duplicate copies are retained and

are well-sampled (Table 1, 2), paralogy can be detected as intra-genome polymorphism shared among the species that have arisen since the duplication occurred. Well-sampled copies can potentially provide valuable phylogenetic information. Sparse sampling when paralogy is unsuspected (Table 1, 3), or random loss of paralogs after duplication (Table 1, 4), will yield similar patterns that are difficult to interpret [136,137], such as widespread taxon polyphyly, sporadic intra-individual polymorphism, and discordance among gene trees. Paralogy does not appear to have major effects in the present analyses, although it is difficult to entirely rule it out [137], especially for the GBSSI data set. Initial paralogy assessments for pepC,  $\beta$ -amylase, and GBSSI were based on information about copy number from crop grasses, combined with preliminary phylogenetic analyses of available sequences in Genbank. The primers for the pepC gene [138] were designed to amplify a single copy from the small pepC gene family [139-141], and similarly, the primers for the  $\beta$ -amylase gene [142] target one copy from a small gene family [143]. Neither tree shows evidence of multiple paralogs in the form of shared intra-individual polymorphism or widespread polyphyly. The

GBSSI tree is more complicated in terms of taxon polyphyly, unanticipated gene copies (especially in *E. repens*), unexpected relationships among diploid taxa, and occasional intra-individual polymorphism (e.g., *P. spicata*). The GBSSI gene was presumed to be a single copy in the wheat tribe [144] based on earlier studies in crop grasses [145–149]. Recent phylogenetic studies in grasses are generally consistent with a single GBSSI copy [150–154], although duplication was inferred in *Spartina* [155]. Based on the general consensus among existing studies, and the lack of the characteristic pattern of shared intra-individual polymorphisms on the Triticeae GBSSI tree that would indicate paralogy, the sequences used to generate the present GBSSI tree are assumed to be orthologs, with the complicated patterns resulting from other phenomena.

Allopolyploid sequences that lack clear affinities to extant diploids could reflect incomplete sampling among diploids, or that the progenitor is extinct or undiscovered (Table 1, 5). The Y genome from the Asian tetraploids is phylogenetically distinct on all three nuclear gene trees (Figures 3, 4, 5a), and represents a case where the diploid progenitor's lineage is probably either extinct or so rare as to have remained undiscovered. The long-standing interest in the wheat tribe has led to extensive collection efforts of wild representatives throughout the world over many decades. Throughout numerous cytogenetic studies [66,68,71–74,76–78,98–101] and subsequent phylogenetic analyses [40,102,133], no diploid Y-genome species has ever been confirmed; the Y genome is a distinct evolutionary lineage known only as a component of certain allopolyploid genomes.

A second case of a “missing diploid” involves the phylogenetically distinct unknown (“UK”) sequence clade from the hexaploid *E. repens*. It is present on all three nuclear gene trees, suggesting that it represents an entire genome of unknown origin, rather than introgression of a limited number of loci. It is widespread within *E. repens*, appearing in individuals from its native Eurasian range [52] and in introduced plants from widely-separated sites in the United States [113]. In contrast to the well-studied Y genome from Asian *Elymus*, the “UK” genome is known only from recent phylogenetic analyses of *E. repens*, so it is more difficult to determine whether the missing diploid is best explained as a sampling artifact vs. an extinct or undiscovered lineage.

If similar polyploid combinations arise recurrently (Table 1, 6), polyploid alleles can be polyphyletic within diploid donor clades. Given several major assumptions – adequate phylogenetic resolution, no prior introgression among donor diploids, no subsequent introgression among polyploids, and no homoeolog loss – each event will potentially result in a separate clade of polyploid sequences arising from within each diploid donor clade. Because entire genomes are involved, recurrent origins should simultaneously affect multiple nuclear gene trees in similar ways. While the design the present study precludes detection of multiple origins from within the same species, recurrent origins of the broader **StStHH** or **StStYY** combinations are potentially detectable, because there is enough variation within *Pseudoroegneria* (**St**) and *Hordeum* (**H**) to provide phylogenetic signal within each clade (Figures 2–5). There are several cases where polyploid homolog groups are

in fact polyphyletic, including the **St** sequences of the **StStYY** species (Figures 3, 4, 5a) and the North American **StStHH** species (Figures 4, 5a); and the **H** sequences of the North American and Eurasian **StStHH** species (Figure 5b) and of *E. repens* (Figures 3, 4, 5b). However, although polyphyly is seen across multiple trees in some of these cases, the specific polyphyletic patterns are not shared among the trees; thus, the patterns look more consistent with multiple independent introgression events than with recurrent **StStHH** or **StStYY** origins (although subsequent gene exchange among the lineages could obscure shared patterns that would indicate multiple origins).

Introgression among species (Table 1, 7–11) can lead to tree-specific effects and among-tree incongruence. Few of the phylogenetic patterns observed here are clearly attributable to introgression within the provisional species groups (Table 1, 7, 8). For example, in the Y-genome clade of the **StStYY** group, within-species samples (*E. abolinii*, *E. antiquus*, *E. caucasicus*, *E. ciliaris*, and *E. semicostatus*) are either monophyletic or consistent with monophyly; gene exchange among **StStYY** species would disrupt this. The **St** sequences from the **StStYY** group are more complicated in that they are broadly polyphyletic on all three trees, but this broad pattern suggests introgression from outside of the **StStYY** group, rather than among species within it. Similarly, placement of the **StStHH** species reflects among-group introgression involving  $\beta$ -amylase (Figure 4) and GBSSI (Figure 5), but there is no clear evidence of within-group gene exchange. While there is too little resolution to address gene exchange among the **StStHH** species in detail, the consistency of certain relationships across trees (*E. canadensis*, *E. hystrix*, *E. riparius*, and *E. virginicus* **H** sequences; *E. brachyaristatus* and *E. sibiricus* **St** sequences) speaks against extensive interspecific gene exchange among the **StStHH** species. *Elymus mutabilis* represents a possible exception; its placement relative to *E. caninus*, *E. dentatus*, and *E. sibiricus* hints at a history of either within-group gene exchange or incomplete lineage sorting.

Introgression from taxa outside of a polyploid group (Table 1, 9–11) will result in polyphyly of a polyploid's homoeologs coupled with incongruence among the gene trees, which is a widespread combination of phenomena here. For example, the **St** sequences of the Asian **StStYY** species are polyphyletic on all three nuclear trees (Figure 3, 4, 5a), most dramatically so on the GBSSI tree (Figure 5a), and the details of the patterns differ among all three trees. In addition, while the pepC tree (Figure 3) points to the Eurasian *P. strigosa* as their potential progenitor, the  $\beta$ -amylase tree (Figure 4) instead places the North American species *P. spicata* as their closest diploid relative (the GBSSI tree is equivocal). Furthermore, the donor genus *Pseudoroegneria* is itself polyphyletic on the GBSSI tree (Figure 5a), and *E. antiquus* shares an anomalous allele with *P. libanotica* and *P. tauri*, whether through a separate polyploidization event or through introgression from *Pseudoroegneria* into *E. antiquus*. The broad polyphyly of these sequences within trees, along with their conflicting placement among trees, highlight the acquisition of alleles from outside the **StStYY** group through introgression. It appears to

be a one-way process, in that no **Y**-genome sequences appear in any other species.

Within the **H**-genome clade, the Eurasian **StStHH** species show little differentiation on the pepC and  $\beta$ -amylase trees (Figures 3, 4), but are widely polyphyletic on the GBSSI tree (Figure 5b). The *E. repens* **H** homoeologs are polyphyletic on all three nuclear trees, even though its chloroplast sequences (Figure 2) and most of its other nuclear homoeologs are either monophyletic or consistent with monophyly. Both of these cases can be explained by multiple instances of introgression from *Hordeum* into *Elymus*. Exchange between *Hordeum* and *Elymus* might, but need not, involve gene exchange across ploidy levels, as there are numerous tetraploid and hexaploid *Hordeum* species [58].

*Elymus repens* not only exhibits introgression from *Hordeum*, but from several other sources. A *Taeniatherum*-like clade of *E. repens* sequences appears only on the GBSSI tree (Figure 5a), and is thus probably a result of introgression. It is found in native accessions [52] as well as the introduced U. S. accessions analyzed here. Furthermore, the placement of *Taeniatherum* is itself anomalous on the GBSSI tree, away from its usual position in or near the *Triticum/Aegilops* group (Figures 2–4), so *Taeniatherum*'s GBSSI copy was itself probably gained through introgression. Whether these two events were independent or sequential is not clear. The more distant Poae-like GBSSI copy (Figure 6), as well as apparent *Bromus*- and *Panicum*-like ITS sequences [52] reveal this species to be a highly complex entity, with a history of acquiring genetic material from distant sources.

Finally, introgression among diploid species (Table 1, 11) can lead to gene-tree incongruence involving both diploids and their derived polyploids. One possible case was mentioned above: *P. tauri* and *P. libanotica* are outside the main **St** sequence group on the GBSSI gene (Figure 5a), along with the **StStYY** species *E. antiquus*. A possible scenario is that *P. tauri* and *P. libanotica* are misplaced due to the acquisition a foreign GBSSI allele, which was then passed to *E. antiquus* through hybridization, or during an independent polyploidization event. Another possible case involves the diphyetic *P. spicata* on the GBSSI tree: one *P. spicata* allele (6c) is found in a largely Eurasian clade with *P. strigosa* and a mix of Eurasian and North American **StStHH** species. Thus, *P. spicata* might have acquired a Eurasian **St** allele through introgression, and passed it into the North American **StStHH** group through hybridization or during polyploidization.

Different phenomena have different effects – on polyploids vs. diploids, and on single gene trees vs. multiple trees – so it is reasonable to favor some processes over others based on gene tree topologies. However, different processes can produce very similar effects, and when they act simultaneously, they can alter or mask one another's effects. Thus, as the discussion above illustrates, the interpretation of trees in terms of individual evolutionary processes is speculative in all but the most straightforward situations.

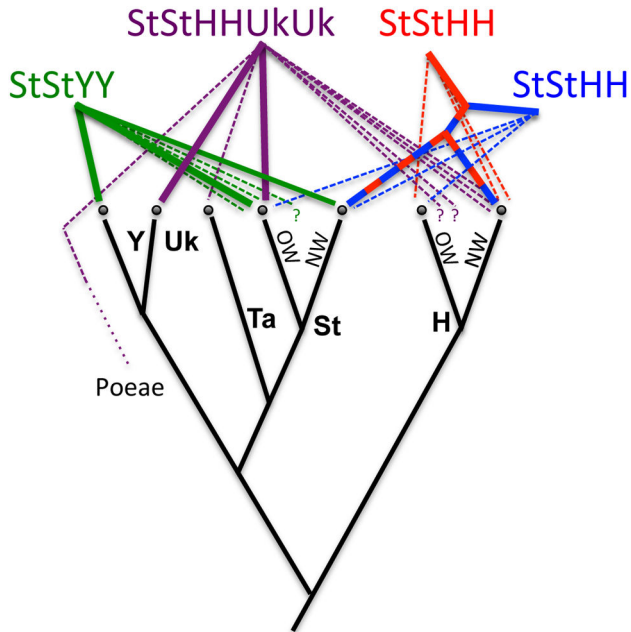
### Polyploid Classification

Classification of stable reticulate taxa creates a logical conundrum in a hierarchical system [156–159], even for the

most straightforward allopolyploid taxa arising from a single origin, and with no subsequent introgression. The recognition of an allotetraploid lineage as a separate named genus makes sense in some ways – from a population standpoint, species that arise through diversification from a single polyploid ancestor can be defined as a monophyletic group (e.g., as applied by [6,22]). On the other hand, recognition of an allopolyploid lineage as a taxon renders both of the originating groups paraphyletic [160,161]. For example, even if the **StStHH** tetraploids arose very simply via a single origin, their recognition as *Elymus* leaves *Pseudoroegneria* and *Hordeum* paraphyletic. As more polyploids are considered (e.g., the **StStYY** group and *E. repens*), the problems with circumscribing *Pseudoroegneria* and *Hordeum* increase. Mindell [162], using holobiont phylogeny as an example of reticulate evolution, recommended expanding the definition of groups from which a composite taxon is derived. Thus, in his example, green plant chloroplasts would be properly classified within the *Cyanobacteria*. Applying this approach to an intergeneric polyploid lineage – separately classifying the **St** and **H** genomes of *Elymus* within their ancestral genera *Pseudoroegneria* and *Hordeum* – has some logical appeal, especially insofar as the **St** and **H** homoeologs do not pair at meiosis, and each genome's integrity is maintained within the *Elymus* polyploid nucleus. This would also open the door to naming the **Y**-genome clade, even though it is only known to exist as a component within larger allopolyploid genomes. Overall, however, this approach to classification is unintuitive and impractical when applied to individual plants and taxa. A simpler approach would be to classify the polyploids and all of their progenitors as a single genus, thus avoiding the entire problem of paraphyletic genera. Ultimately, this would result in defining most or all of the tribe as a single genus, which is not without precedent [163], although this would do little more than push the same classification issues to the subgeneric level. Furthermore, while a single-genus solution might reasonably be proposed for a more obscure group, it would be unlikely to gain acceptance for this tribe which, due to its economic importance, has undergone a long history of close scrutiny [164] and intense taxonomic focus [165].

### Conclusion

Ultimately, the results and discussion presented here might have done less to clarify the history of *Elymus* than to highlight the difficulty of developing a step-by-step, gene-by-gene reconstruction of the evolutionary events leading to the present situation (Figure 8). This is appropriate to the final impression I hope to leave – that of a genetically tangled set of lineages that are evolutionarily distinct, but are neither entirely separable from one another, nor fully cohesive individually.



**Figure 8. Summary diagram of *Elymus* origins.** Unbroken lines represent major contributions, and dotted lines show minor contributions. Colors follow Figures 1–7. The simple bifurcating division of *Pseudoroegneria* (St) and *Hordeum* (H) progenitor species into Old World (OW) and New World (NW) lineages is used for illustrative purposes and does not reflect their true phylogenetic complexity, particularly within *Hordeum*.  
doi: 10.1371/journal.pone.0078449.g008

## Supporting Information

**Dataset S1. Chloroplast trnT/L/F genes and spacers, and rpoA gene sequences (see Figure 2).**

(TXT)

**Dataset S2. Phosphoenolpyruvate carboxylase partial gene sequences (see Figure 3).**

(TXT)

**Dataset S3. Beta-amylase partial gene sequences (see Figure 4).**

(TXT)

**Dataset S4. Granule-bound starch synthase partial gene sequences (see Figure 5).**

(TXT)

**Dataset S5. Granule-bound starch synthase partial gene sequences with introns excluded (see Figure 6).**

(TXT)

**Table S1. List of taxa, collection information, and Genbank accession numbers for sequences analyzed for Figures 2–5.**

(PDF)

**Table S2. Sequences used in GBSSI exon analysis (Figure 6).**

(PDF)

**Table S3. Datasets, partitions where applicable, models selected under AIC criterion, and parameters estimated on the best ML tree.**

(PDF)

## Acknowledgements

Thanks to Mary Barkworth and an anonymous reviewer for their suggestions for improving the manuscript; to Melissa Burns and Marianna Naum for assistance with chloroplast DNA data collection; and to the USDA-ARS National germplasm System for seeds.

## Author Contributions

Conceived and designed the experiments: RJMG. Performed the experiments: RJMG. Analyzed the data: RJMG. Contributed reagents/materials/analysis tools: RJMG. Wrote the manuscript: RJMG.

## References

- Popp M, Oxelman B (2001) Inferring the history of the polyploid *Silene aegaea* (Caryophyllaceae) using plastid and homoeologous nuclear DNA sequences. *Mol Phylogenet Evol* 20: 474–481. doi:10.1006/mpev.2001.0977. PubMed: 11527472.
- Hughes CE, Bailey CD, Harris SA (2002) Divergent and reticulate species relationships in *Leucaena* (Fabaceae) inferred from multiple data sources: insights into polyploid origins and nrDNA polymorphism. *Am J Bot* 89: 1057–1073. doi:10.3732/ajb.89.7.1057. PubMed: 21665706.
- Drábková LZ, Kirschner J, Štěpánek J, Závěský L, Vlček Č (2009) Analysis of nrDNA polymorphism in closely related diploid sexual, tetraploid sexual, and polyploid agamous species. *Plant Syst Evol* 278: 67–85. doi:10.1007/s00606-008-0134-8.
- Sang T, Crawford DJ, Stuessy TF (1995) Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. *Proc Natl Acad Sci U S A* 92: 6813–6817. doi:10.1073/pnas.92.15.6813. PubMed: 7624325.
- Alvarez I, Wendel JF (2003) Ribosomal ITS sequences and plant phylogenetic inference. *Mol Phylogenet Evol* 29: 417–434. doi:10.1016/S1055-7903(03)00208-2. PubMed: 14615184.
- Grover CE, Grupp KK, Wanzek RJ, Wendel JF (2012) Assessing the monophyly of polyploid *Gossypium* species. *Plant Syst Evol* 298: 1177–1183. doi:10.1007/s00606-012-0615-7.
- Ainouche ML, Baume A, Salmon A (2004) *Spartina anglica* CE Hubbard: a natural model system for analyzing early evolutionary changes that affect allopolyploid genomes. *Biol J Linn Soc* 82: 475–484. doi:10.1111/j.1095-8312.2004.00334.x.
- Brysting AK, Oxelman B, Huber KT, Moulton V, Brochmann C (2007) Untangling complex histories of genome mergings in high polyploids. *Syst Biol* 56: 467–476. doi:10.1080/10635150701424553. PubMed: 17562470.
- Emshwiller E, Doyle JJ (2002) Origins of domestication and polyploidy in oca (*Oxalis tuberosa*: Oxalidaceae). 2. Chloroplast-expressed glutamine synthase data. *Am J Bot* 89: 1042–1056. doi:10.3732/ajb.89.7.1042. PubMed: 21665705.
- Joly S, Starr JR, Lewis WH, Bruneau A (2006) Polyploid and hybrid evolution in roses east of the Rocky Mountains. *Am J Bot* 93: 412–425. doi:10.3732/ajb.93.3.412. PubMed: 21646201.
- Lihová J, Shimizu KK, Marhold K (2006) Allopolyploid origin of *Cardamine asarifolia* (Brassicaceae): incongruence between plastid and nuclear ribosomal DNA sequences solved by a single-copy nuclear gene. *Mol Phylogenet Evol* 39: 759–786. doi:10.1016/j.ympev.2006.01.027. PubMed: 16527494.
- 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. doi:10.1016/j.ympev.2006.01.023. PubMed: 16504543.
- Popp M, Erixon P, Eggens F, Oxelman B (2005) Origin and evolution of a circumpolar polyploid species complex in *Silene* (Caryophyllaceae) inferred from low copy nuclear RNA polymerase introns, rDNA, and chloroplast DNA. *Syst Bot* 30: 302–313. doi:10.1600/0363644054223648.
- Rodríguez F, Spooner DM (2009) Nitrate reductase phylogeny of potato (*Solanum* sect. *Petota*) genomes with emphasis on the origins of the polyploid species. *Syst Bot* 34: 207–219. doi:10.1600/036364409787602195.
- Smedmark JEE, Eriksson T, Bremer B (2005) Allopolyploid evolution in Geinae (Colurieae : Rosaceae) - building reticulate species trees from bifurcating gene trees. *Organ Divers Evol* 5: 275–283. doi:10.1016/j.ode.2004.12.003.
- Straub SCK, Pfeil BE, Doyle JJ (2006) Testing the polyploid past of soybean using a low-copy nuclear gene – Is *Glycine* (Fabaceae : Papilionoideae) an auto- or allopolyploid? *Mol Phylogenet Evol* 39: 580–584. doi:10.1016/j.ympev.2005.11.018. PubMed: 16384718.
- Johnson LA, Chan LM, Pozner R, Glazier LD (2012) Allotetraploids in Patagonia with affinities to western North American diploids: did dispersal or genome doubling occur first? *Bot Rev* 78: 288–306. doi:10.1007/s12229-012-9101-8.
- Cai D, Rodriguez F, Teng Y, Ané C, Bonierbale M et al. (2012) Single copy nuclear gene analysis of polyploidy in wild potatoes (*Solanum* section *Petota*). *BMC Evol Biol* 12: 70. doi:10.1186/1471-2148-12-70. PubMed: 22624678.
- Bendiksby M, Tribsch A, Borgen L, Trávníček P, Brysting AK (2011) Allopolyploid origins of the *Galeopsis* tetraploids – revisiting Müntzing's classical textbook example using molecular tools. *New Phytol* 191: 1150–1167. doi:10.1111/j.1469-8137.2011.03753.x. PubMed: 21599670.



20. Brokaw JM, Hufford L (2010) Origins and introgression of polyploid species in *Mentzelia* section *Trachyphytum* (Loasaceae). *Am J Bot* 97: 1457–1473. doi:10.3732/ajb.0900388. PubMed: 21616900.
21. Triplett JK, Wang Y, Zhong J, Kellogg EA (2012) Five nuclear loci resolve the polyploid history of switchgrass (*Panicum virgatum*, p. L.) and relatives. *PLoS ONE* 7: e38702.
22. Marcussen T, Jakobsen KS, Danihelka J, Ballard HE, Blaxland K et al. (2012) Inferring species networks from gene trees in high-polyploid North American and Hawaiian violets (*Viola*, Violaceae). *Syst Biol* 61: 107–126. doi:10.1093/sysbio/syr096. PubMed: 21918178.
23. Sessa EB, Zimmer EA, Givnish TJ (2012) Reticulate evolution on a global scale: a nuclear phylogeny for New World *Dryopteris* (Dryopteridaceae). *Mol Phylogenet Evol* 64: 563–581. doi:10.1016/j.ympev.2012.05.009. PubMed: 22634937.
24. Fan X, Sha L-N, Zeng J, Kang H-Y, Zhang H-Q et al. (2012) Evolutionary dynamics of the *Pgk1* gene in the polyploid genus *Kengyilia* (Triticeae: Poaceae) and its diploid relatives. *PLOS ONE* 7: e31122. doi:10.1371/journal.pone.0031122. PubMed: 22363562.
25. Buggs RJA, Doust AN, Tate JA, Koh J, Soltis K et al. (2009) Gene loss and silencing in *Tragopogon miscellus* (Asteraceae): comparison of natural and synthetic allotetraploids. *Heredity* 103: 73–81. doi:10.1038/hdy.2009.24. PubMed: 19277058.
26. Kashkush K, Feldman M, Levy AA (2002) Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. *Genetics* 160: 1651–1659. PubMed: 11973318.
27. Song KM, Lu P, Tang K, Osborn TC (1995) Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. *Proc Natl Acad Sci U S A* 92: 7719–7723. doi:10.1073/pnas.92.17.7719. PubMed: 7644483.
28. Town CT, Cheung F, Maiti R, Crabtree J, Haas BJ et al. (2006) Comparative genomics of *Brassica oleracea* and *Arabidopsis thaliana* reveal gene loss, fragmentation, and dispersal after polyploidy. *Plant Cell* 18: 1348–1359. doi:10.1105/tpc.106.041665. PubMed: 16632643.
29. Leitch IJ, Bennett MD (2004) Genome downsizing in polyploid plants. *Biol J Linn Soc* 82: 651–663. doi:10.1111/j.1095-8312.2004.00349.x.
30. Smedmark JEE, Eriksson T, Evans RC, Campbell CS (2003) Ancient allopolyploid speciation in Geinae (Rosaceae): evidence from nuclear granule-bound starch synthase (GBSSI) gene sequences. *Syst Biol* 52: 374–385. doi:10.1080/10635150309332. PubMed: 12775526.
31. Smith JF, Funke MM, Woo VL (2006) A duplication of *gcyc* predates divergence within tribe Coronanthereae (Gesneriaceae): phylogenetic analysis and evolution. *Plant Syst Evol* 261: 245–256. doi:10.1007/s00606-006-0445-6.
32. Sang T, Zhang D (1999) Reconstructing hybrid speciation using sequences of low copy nuclear genes: hybrid origins of five *Paeonia* species based on *Adh* gene phylogenies. *Syst Bot* 24: 148–163. doi:10.2307/2419546.
33. Cronn RC, Small RL, Wendel JF (1999) Duplicated genes evolve independently after polyploid formation in cotton. *Proc Natl Acad Sci U S A* 96: 14406–14411. doi:10.1073/pnas.96.25.14406. PubMed: 10588718.
34. Schranz ME, Mitchell-Olds T (2006) Independent ancient polyploidy events in the sister families Brassicaceae and Cleomaceae. *Plant Cell* 18: 1152–1165. doi:10.1105/tpc.106.041111. PubMed: 16617098.
35. Marcussen T, Oxelman B, Skog A, Jakobsen KS (2010) Evolution of plant RNA polymerase IV/5 genes: evidence of subneofunctionalization of duplicated *NRPD2/NRPE2*-like paralogs in *Viola* (Violaceae). *BMC Evol Biol* 10: 45. doi:10.1186/1471-2148-10-45. PubMed: 20158916.
36. Ferguson D, Sang T (2001) Speciation through homoploid hybridization between allotetraploids in peonies (*Paeonia*). *Proc Natl Acad Sci U S A* 98: 3915–3919. doi:10.1073/pnas.061288698. PubMed: 11259655.
37. Winkworth RC, Donoghue MJ (2004) *Viburnum* phylogeny: evidence from the duplicated nuclear gene GBSSI. *Mol Phylogenet Evol* 33: 109–126. doi:10.1016/j.ympev.2004.05.006. PubMed: 15324842.
38. Brassac J, Jakob SS, Blattner FR (2012) Progenitor-derivative relationships of *Hordeum* polyploids (Poaceae, Triticeae) inferred from sequences of *TOPO6*, a nuclear low-copy gene. *PLOS ONE* 3: e33808.
39. Lihová J, Marhold K, Kudoh H, Koch MA (2006) Worldwide phylogeny and biogeography of *Cardamine flexuosa* (Brassicaceae) and its relatives. *Am J Bot* 93: 1206–1221. doi:10.3732/ajb.93.8.1206. PubMed: 21642185.
40. Mason-Gamer RJ, Burns MM, Naum M (2010) Phylogenetic relationships and reticulation among Asian *Elymus* (Poaceae) allotetraploids: analysis of three nuclear genes. *Mol Phylogenet Evol* 54: 10–22. doi:10.1016/j.ympev.2009.10.002. PubMed: 1981864.
41. Roelofs D, van Velzen J, Kuperus P, Bachmann K (1997) Molecular evidence for an extinct parent of the tetraploid species *Microseris acuminata* and *M. campestris* (Asteraceae, Lactuceae). *Mol Ecol* 6: 641–649. doi:10.1046/j.1365-294X.1997.00225.x. PubMed: 9226946.
42. Ainouche ML, Baumel A, Salmon A, Yannic G (2003) Hybridization, polyploidy and speciation in *Spartina* (Poaceae). *New Phytol* 161: 165–172. doi:10.1046/j.1469-8137.2003.00926.x.
43. Arft AM, Ranker TA (1998) Allopolyploid origin and population genetics of the rare orchid *Spiranthes diluvialis*. *Am J Bot* 85: 110–122. doi:10.2307/2446560. PubMed: 21684885.
44. Cook LM, Soltis PS, Brunsfield SJ, Soltis DE (1998) Multiple independent formations of *Tragopogon* tetraploids (Asteraceae): evidence from RAPD markers. *Mol Ecol* 7: 1293–1302. doi:10.1046/j.1365-294x.1998.00453.x.
45. Soltis DE, Soltis PS (1999) Polyploidy: recurrent formation and genome evolution. *Trends Ecol Evol* 14: 348–352. doi:10.1016/S0169-5347(99)01638-9. PubMed: 10441308.
46. Church SA, Taylor DR (2005) Speciation and hybridization among *Houstonia* (Rubiaceae) species: The influence of polyploidy on reticulate evolution. *Am J Bot* 92: 1372–1380. doi:10.3732/ajb.92.8.1372. PubMed: 21646157.
47. Garcia-Jacas N, Soltis PS, Font M, Soltis DE, Vilatersana R et al. (2009) The polyploid series of *Centaureia toletana*: glacial migration and introgression revealed by nrDNA and cpDNA sequence analyses. *Mol Phylogenet Evol* 52: 377–394. doi:10.1016/j.ympev.2009.03.010. PubMed: 19306936.
48. Hedrén M, Nordström S, Bateman RM (2011) Plastid and nuclear DNA marker data support the recognition of four tetraploid marsh orchids (*Dactylorhiza majalis* s.l., Orchidaceae) in Britain and Ireland, but require their recircumscription. *Biol J Linn Soc* 104: 107–128. doi:10.1111/j.1095-8312.2011.01708.x.
49. Jørgensen MH, Ehrich D, Schmickl R, Koch MA, Brysting AK (2011) Interspecific and interploidal gene flow in central European *Arabidopsis* (Brassicaceae). *BMC Evol Biol* 11: 346. doi:10.1186/1471-2148-11-346. PubMed: 22126410.
50. Mahelka V, Fehrer J, Krahulec F, Jarolímová V (2007) Recent natural hybridization between two allopolyploid wheatgrasses (*Elytrigia*, Poaceae): ecological and evolutionary implications. *Ann Bot* 100: 249–260. doi:10.1093/aob/mcm093. PubMed: 17562679.
51. Doyle JJ, Doyle JL, Rauscher JT, Brown AHD (2004) Diploid and polyploid reticulate evolution throughout the history of the perennial soybeans (*Glycine* subgenus *Glycine*). *New Phytol* 161: 121–132.
52. Mahelka V, Kopecký D (2010) Gene capture from across the grass family in *Elymus repens* (L.). Gould (Poaceae, Triticeae) as evidenced by ITS, GBSSI, and molecular cytogenetics. *Molecular Biology and Evolution* 27: 1370–1390.
53. Chapman MA, Abbott RJ (2010) Introgression of fitness genes across a ploidy barrier. *New Phytol* 186: 63–71. doi:10.1111/j.1469-8137.2009.03091.x. PubMed: 19912548.
54. Ma J-X, Li N-Y, Vogl C, Ehrendorfer F, Guo Y-P (2010) Allopolyploid speciation and ongoing backcrossing between diploid progenitor and tetraploid progeny lineages in the *Achillea millefolium* species complex: analyses of single-copy nuclear genes and genomic AFLP. *BMC Evol Biol* 10: 100. doi:10.1186/1471-2148-10-100. PubMed: 20388203.
55. Oberle B, Montgomery RA, Beck JB, Esselman E (2012) A morphologically intergrading population facilitates plastid introgression from diploid to tetraploid *Dodecatheon* (Primulaceae). *Bot J Linn Soc* 168: 91–100. doi:10.1111/j.1095-8339.2011.01191.x.
56. Thórsson AET, Pálsson S, Lascoux M, Anamthawat-Jónsson K (2010) Introgression and phylogeography of *Betula nana* (diploid), *B. pubescens* (tetraploid) and their triploid hybrids in Iceland inferred from cpDNA haplotype variation. *J Biogeogr* 37: 2098–2110. doi:10.1111/j.1365-2699.2010.02353.x.
57. Laureto PJ, Barkman TJ (2011) Nuclear and chloroplast DNA suggest a complex single origin for threatened allopolyploid *Solidago houghtonii* (Asteraceae) involving reticulate evolution and introgression. *Syst Bot* 36: 209–226. doi:10.1600/036364411X553289.
58. Löve Å (1984) *Conspectus of the Triticeae*. *Feddes Repertorium* 95: 425–521.
59. Flora of China Editorial Committee (2006) *Flora of China*, Volume 22: Poaceae. Beijing, St. Louis: Science Press, Missouri Botanical Garden Press.
60. Flora of North America Editorial Committee (2007) *Flora of North America*, Volume 24: Commelinidae (in part): Poaceae, part I. Oxford: Oxford University Press.
61. Tzvelev NN (1984) *Grasses of the Soviet Union*. BR Sharma. Rotterdam: A. A. Balkema Publishing House.
62. Webb DA (1980) *Flora Europaea*. Volume 5. Alismataceae to Orchidaceae. Cambridge University Press.
63. Dewey DR (1967) Synthetic hybrids of new world and old world *Agropyron*s: III. *Agropyron repens* X tetraploid *Agropyron spicatum*. *Am J Bot* 54: 93–98. doi:10.2307/2440890.

64. Dewey DR (1968) Synthetic *Agropyron-Elymus* hybrids: III. *Elymus canadensis* X *Agropyron caninum*, *A. trachycaulum*, and *A. striatum*. American Journal of Botany 55: 1133–1139.
65. Dewey DR (1970) Hybrids and induced amphiploids of *Agropyron dasystachyum* X *Agropyron caninum*. Bot Gaz 131: 342–348. doi: 10.1086/336551.
66. Dewey DR (1974) Cytogenetics of *Elymus sibiricus* and its hybrids with *Agropyron tauri*, *Elymus canadensis*, and *Agropyron caninum*. Bot Gaz 135: 80–87. doi:10.1086/337824.
67. Dewey DR (1984) The genomic system of classification as a guide to intergeneric hybridization within the perennial Triticeae. In: JP Gustafson. Gene manipulation in plant improvement, Proc 16th Stadler Genetics Symposium: New York, NY. Plenum Publishing Company. pp. 209–279.
68. Jensen KB (1990) Cytology and morphology of *Elymus pendulinus*, *E. pendulinus* ssp. *multiculmis*, and *E. parviglume* (Poaceae: Triticeae). Bot Gaz 151: 245–251. doi:10.1086/337824.
69. Jensen KB (1993) Cytogenetics of *Elymus magellanicus* and its intra- and inter-generic hybrids with *Pseudoroegneria spicata*, *Hordeum violaceum*, *E. trachycaulum*, *E. lanceolatus*, and *E. glaucus* (Poaceae: Triticeae). Genome 36: 72–76.
70. Jensen KB (1996) Genome analysis of Eurasian *Elymus thordianus*, *E. melantherus*, and *E. kokonoricus* (Triticeae: Poaceae). Int J Plant Sci 157: 136–141. doi:10.1086/297330.
71. Jensen KB, Hatch SL (1989) Genome analysis, morphology, and taxonomy of *Elymus gmelinii* and *E. strictus* (Poaceae: Triticeae). Bot Gaz 150: 84–92. doi:10.1086/337752.
72. Lu B-R (1993) Genomic relationships within the *Elymus parviglumis* group (Triticeae: Poaceae). Plant Syst Evol 187: 191–211. doi:10.1007/BF00994099.
73. Lu B-R, Salomon B (1992) Differentiation of the **SY** genomes in Asiatic *Elymus*. Hereditas 116: 121–126.
74. Lu B-R, Salomon B (2004) Differentiation of the **StY** genomes in *Elymus* species as referred by meiotic pairing in interspecific hybrids and its evolutionary significance. Biodivers Sci 12: 213–226.
75. Lu B-R, Salomon B, von Bothmer R (1991) Meiotic studies of the hybrids among *Pseudoroegneria cognata*, *Elymus semicostatus*, and *E. pendulinus* (Poaceae). Hereditas 114: 117–124.
76. Salomon B (1993) Interspecific hybridizations in the *Elymus semicostatus* group (Poaceae). Genome 36: 889–905. PubMed: 18470036.
77. Salomon B, Lu B-R (1992) Genomic groups, morphology, and sectional delimitation in Eurasian *Elymus* (Poaceae, Triticeae). Plant Syst Evol 180: 1–13. doi:10.1007/BF00940394.
78. Salomon B, Lu B-R (1994) Interspecific hybridizations among species of the *Elymus semicostatus* and *Elymus tibeticus* groups (Poaceae). Plant Syst Evol 189: 1–13. doi:10.1007/BF00937574.
79. Barkworth ME, Cutler DR, Rollo JS, Jacobs SWL, Rashid A (2009) Morphological identification of genomic genera in the Triticeae. Breeding Sci 59: 561–570. doi:10.1270/jsbbs.59.561.
80. Barkworth ME, Jacobs SWL (2011) The Triticeae (Gramineae) in Australasia. Telopea 13: 37–56.
81. Yen C, Yang JL, Baum BR (2005) Hitoshi Kihara, Åskell Löve and the modern genetic concept of the genera in the tribe Triticeae (Poaceae). Acta Phytotaxonomica Sin 43: 82–93. doi:10.1360/aps040073.
82. Yen C, Yang JL (2009) Historical review and prospect of taxonomy of tribe Triticeae Dumortier (Poaceae). Breeding Sci 59: 513–518. doi: 10.1270/jsbbs.59.513.
83. Dewey DR (1975) The origin of *Agropyron smithii*. Am J Bot 62: 524–530. doi:10.2307/2441961.
84. Chen Q, Conner RL, Laroche A, Thomas JB (1998) Genome analysis of *Thinopyrum intermedium* and *Thinopyrum ponticum* using genomic in situ hybridization. Genome 41: 580–586. doi:10.1139/g98-055. PubMed: 9796107.
85. Liu Z-W, Wang RR-C (1993) Genome analysis of *Elytrigia caespitosa*, *Lophopyrum nodosum*, *Pseudoroegneria geniculata* ssp. *scythia*, and *Thinopyrum intermedium* (Triticeae, Gramineae). Genome 36: 102–111. doi:10.1139/g93-014. PubMed: 18469974.
86. Zhang XY, Dong YS, Wang RR-C (1996) Characterization of genomes and chromosomes in partial amphiploids of the hybrid *Triticum aestivum* x *Thinopyrum ponticum* by in situ hybridization, isozyme analysis, and RAPD. Genome 39: 1062–1071. doi:10.1139/g96-133. PubMed: 18469955.
87. Arterburn M, Kleinhofs A, Murray T, Jones S (2011) Polymorphic nuclear gene sequences indicate a novel genome donor in the polyploid genus *Thinopyrum*. Hereditas 148: 8–27. doi:10.1111/j.1601-5223.2010.02084.x. PubMed: 21410467.
88. Mahelka V, Kopecký D, Paštová L (2011) On the genome constitution and evolution of intermediate wheatgrass (*Thinopyrum intermedium*: Poaceae, Triticeae). BMC Evol Biol 11: 127. doi: 10.1186/1471-2148-11-127. PubMed: 21592357.
89. Church GL (1958) Artificial hybrids of *Elymus virginicus* with *E. canadensis*, *interruptus*, *riparius*, and *wiegandii*. Am J Bot 45: 410–417. doi:10.2307/2439642.
90. Dewey DR (1982) Genomic and phylogenetic relationships among North American perennial Triticeae. In: JR Estes. Grasses and grasslands. Norman: Oklahoma University Press. pp. 51–81.
91. Dewey DR (1983) New nomenclatural combinations in the North American perennial Triticeae (Gramineae). Brittonia 35: 30–33. doi: 10.2307/2806044.
92. Dewey DR (1970) Genome relations among *Elymus canadensis*, *Elymus triticoides*, *Elymus dasystachyus*, and *Agropyron smithii*. Am J Bot 57: 861–866. doi:10.2307/2441345.
93. Mason-Gamer RJ (2001) Origin of North American species of *Elymus* (Poaceae: Triticeae) allotetraploids based on granule-bound starch synthase gene sequences. Syst Bot 26: 757–768.
94. Mason-Gamer RJ, Burns MM, Naum M (2010) Reticulate evolutionary history of a complex group of grasses: phylogeny of *Elymus StStHH* allotetraploids based on three nuclear genes. PLOS ONE 5: e10989. doi:10.1371/journal.pone.0010989. PubMed: 20543952.
95. Redinbaugh MG, Jones TA, Zhang Y (2000) Ubiquity of the **St** chloroplast genome in **St**-containing Triticeae polyploids. Genome 43: 836–852. doi:10.1139/gen-43-5-836. PubMed: 11081974.
96. Sun G, Zhang X (2011) Origin of the **H** genome in **StH**-genomic *Elymus* species based on the single-copy nuclear gene *DMC1*. Genome 54: 655–662. doi:10.1139/g11-036. PubMed: 21848405.
97. Sun G-L, Ni Y, Daley T (2008) Molecular phylogeny of *RPB2* gene reveals multiple origin, geographic distribution of **H** genome, and the relationship of the **Y** genome to other genomes of *Elymus* species. Mol Phylogenet Evol 46: 897–907. doi:10.1016/j.ympev.2007.12.024. PubMed: 18262439.
98. Jensen KB (1989) Cytology, fertility, and origin of *Elymus abolinii* (Drob.) Tzvelev and its hybrids with *Pseudoroegneria spicata*, *E. lanceolatus*, *E. dentatus* ssp. *ugamicus*, and *E. drobovii*. Genome 32: 468–474.
99. Lu B-R, von Bothmer R (1990) Intergeneric hybridization between *Hordeum* and Asiatic *Elymus*. Hereditas 112: 109–116.
100. Lu B-R, von Bothmer R (1993) Meiotic analysis of *Elymus caucasicus*, *E. longearistatus*, and their interspecific hybrids with twenty-three *Elymus* species (Triticeae, Poaceae). Plant Syst Evol 185: 35–53. doi: 10.1007/BF00937719.
101. Lu B-R, Yan J, Yang J-L, Flink J (1990) Biosystematic studies among *Roegneria pendulina* Nevski, *R. ciliaris* (Trin.) Nevski, and *R. kamoji* Owhi in the tribe Triticeae, Gramineae. Acta Bot Yunnanica 12: 161–171.
102. Petersen G, Seberg O, Salomon B (2011) The origin of the **H**, **St**, **W**, and **Y** genomes in allotetraploid species of *Elymus* L. and *Stenostachys Turcz.* (Poaceae: Triticeae). Plant Systematics and Evolution 291: 197–210.
103. Yan C, Sun G, Sun D (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. doi: 10.1371/journal.pone.0026853. PubMed: 22046383.
104. Batcher MS (2002) Element stewardship abstract for *Elytrigia repens* (L.) Desv. ex B.D. Jackson. Arlington, VA: The Nature Conservancy..
105. Cauderon Y (1958) Étude cytogénétique des *Agropyrum* français et de leurs hybrides avec les blés. Ann Amélioration Plantes 8: 389–567.
106. Cauderon Y, Saigne B (1961) New interspecific and intergeneric hybrids involving *Agropyron*. Wheat Info Service 12: 13–14.
107. Dewey DR (1961) Hybrids between *Agropyron repens* and *Agropyron desertorum*. J Hered 52: 13–21.
108. Dewey DR (1970) A cytogenetic study of *Agropyron stipifolium* and its hybrids with *Agropyron repens*. Bulletin of The Torrey Botanical Club 97: 315–320.
109. Dewey DR (1976) Derivation of a new forage grass from *Agropyron repens* X *Agropyron spicatum* hybrids. Crop Sci 16: 175–180. doi: 10.2135/cropsci1976.0011183X001600020003x.
110. Assadi M, Runemark H (1995) Hybridisation, genomic constitution and generic delimitation in *Elymus* s.l. (Poaceae: Triticeae). Plant Syst Evol 194: 189–205. doi:10.1007/BF00982855.
111. Orgaard M, Anamthawat-Jónsson K (2001) Genome discrimination by *in situ* hybridization in Icelandic species of *Elymus* and *Elytrigia* (Poaceae:Triticeae). Genome 44: 275–283. doi:10.1139/gen-44-2-275. PubMed: 11341738.
112. Mason-Gamer RJ (2004) Reticulate evolution, introgression, and intertribal gene capture in an allohexaploid grass. Syst Biol 53: 25–37. doi:10.1080/10635150490424402. PubMed: 14965898.

113. Mason-Gamer RJ (2008) Allohexaploidy, introgression, and the complex phylogenetic history of *Elymus repens* (Poaceae). *Mol Phylogenet Evol* 47: 598–611. doi:10.1016/j.ympev.2008.02.008. PubMed: 18372193.
114. Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol Biol* 17: 1105–1109. doi:10.1007/BF00037152. PubMed: 1932684.
115. Petersen G, Seberg O (1997) Phylogenetic analysis of the Triticeae (Poaceae) based on *rpoA* sequence data. *Mol Phylogenet Evol* 7: 217–230. doi:10.1006/mpev.1996.0389. PubMed: 9126564.
116. Mason-Gamer RJ, Orme NL, Anderson CM (2002) Phylogenetic analysis of North American *Elymus* and the monogenomic Triticeae (Poaceae) using three chloroplast DNA data sets. *Genome* 45: 991–1002. doi:10.1139/g02-065. PubMed: 12502243.
117. Posada D (2008) jModeltest: phylogenetic model averaging. *Mol Biol Evol* 25: 1253–1256. doi:10.1093/molbev/msn083. PubMed: 18397919.
118. Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: NH Munro. *Mammalian protein metabolism*. New York: Academic Press. pp. 21–132.
119. Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 16: 111–120. doi:10.1007/BF01731581. PubMed: 7463489.
120. Zharkikh A (1994) Estimation of evolutionary distances between nucleotide sequences. *J Mol Evol* 39: 315–329. doi:10.1007/BF00160155. PubMed: 7932793.
121. Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 17: 368–376. doi:10.1007/BF01734359. PubMed: 7288891.
122. Hasegawa M, Kishino H, Yano T (1985) Dating the human-ape split by a molecular clock of mitochondrial DNA. *J Mol Evol* 22: 160–170. doi:10.1007/BF02101694. PubMed: 3934395.
123. Tavaré S (1986) Some probabilistic and statistical problems in the analysis of DNA sequences. *Lect Math Life Sci* 17: 57–86.
124. Yang Z (1993) Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. *Mol Biol Evol* 10: 1396–1401. PubMed: 8277861.
125. Gu X, Fu Y-X, Li W-H (1995) Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. *Mol Biol Evol* 12: 546–557. PubMed: 7659011.
126. Waddell PJ, Penny D (1996) Evolutionary trees of apes and humans from DNA sequences. In: AJ Locke CR Peters. *Handbook of symbolic evolution*. Oxford: Clarendon Press. pp. 53–73.
127. Glez-Peña D, Gómez-Blanco D, Reboiro-Jato M, Fdez-Riverola F, Posada D (2010) ALTER: program-oriented format conversion of DNA and protein alignments. *Nucleic Acids Res* 38: W14–W18. doi:10.1093/nar/gkp932. PubMed: 20439312.
128. Akaike H (1974) A new look at the statistical model estimation. *IEEE Trans Autom Contr* 19: 716–723. doi:10.1109/TAC.1974.1100705.
129. Zwickl DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion PhD Thesis Austin: the University of Texas.
130. Jaaska V (1992) Isoenzyme variation in the grass genus *Elymus* (Poaceae). *Hereditas* 117: 11–22.
131. Linde-Laursen I, Seberg O, Salomon B (1994) Comparison of the Giemsa C-banded and N-banded karyotypes of two *Elymus* species, *E. dentatus* and *E. glaucescens* (Poaceae: Triticeae). *Plant Syst Evol* 192: 165–176. doi:10.1007/BF00986250.
132. Sun G, Salomon B (2009) Molecular evolution and origin of tetraploid *Elymus* species. *Breeding Sci* 59: 487–491. doi:10.1270/jsbbs.59.487.
133. Sun G, Komatsuda T (2010) Origin of the Y genome in *Elymus* and its relationship to other genomes in Triticeae based on evidence from elongation factor G (*Ef-G*) gene sequences. *Mol Phylogenet Evol* 56: 727–733. doi:10.1016/j.ympev.2010.03.037. PubMed: 20363342.
134. Okito P, Mott IW, Wu Y, Wang RR-C (2009) A Y-genome specific STS marker in *Pseudoroegneria* and *Elymus* species (Triticeae: Gramineae). *Genome* 52: 391–400. doi:10.1139/G09-015. PubMed: 19370094.
135. Liu Q, Ge S, Tang H, Zhang X, Zhu G et al. (2006) Phylogenetic relationships in *Elymus* (Poaceae: Triticeae) based on nuclear ribosomal internal transcribed spacer and chloroplast *trnL-F* sequences. *New Phytol* 170: 411–420. doi:10.1111/j.1469-8137.2006.01665.x. PubMed: 16608465.
136. Bailey CD, Carr TG, Harris SA, Hughes CE (2003) Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. *Mol Phylogenet Evol* 29: 435–455. doi:10.1016/j.ympev.2003.08.021. PubMed: 14615185.
137. Martin AP, Burg TM (2002) Perils of paralogy: Using HSP70 genes for inferring organismal phylogenies. *Syst Biol* 51: 570–587. doi:10.1080/10635150290069995. PubMed: 12228000.
138. Helfgott DM, Mason-Gamer RJ (2004) The evolution of North American *Elymus* (Triticeae, Poaceae) allotetraploids: evidence from phosphoenolpyruvate carboxylase gene sequences. *Syst Bot* 29: 850–861. doi:10.1600/0363644042451017.
139. Crétin C, Santi S, Keryer E, Lepiniec L, Tagu D et al. (1991) The phosphoenolpyruvate carboxylase gene family in *Sorghum*: promoter structure, amino acid sequences and expression of genes. *Gene* 99: 87–94. doi:10.1016/0378-1119(91)90037-C. PubMed: 2022326.
140. Lepiniec L, Keryer E, Philippe H, Gadal P, Crétin C (1993) *Sorghum* phosphoenolpyruvate carboxylase gene family: structure, function, and molecular evolution. *Plant Mol Biol* 21: 487–502. doi:10.1007/BF00228806. PubMed: 8443342.
141. Lepiniec L, Vidal J, Chollet R, Gadal P, Crétin C (1994) Phosphoenolpyruvate carboxylase: structure, regulation, and evolution. *Plant Sci* 99: 111–124. doi:10.1016/0168-9452(94)90168-6.
142. Mason-Gamer RJ (2005) The b-amylase genes of grasses and a phylogenetic analysis of the Triticeae (Poaceae). *Am J Bot* 92: 1045–1058. doi:10.3732/ajb.92.6.1045. PubMed: 21652489.
143. Ziegler P (1999) Cereal beta-amylases. *J Cereal Sci* 29: 195–204. doi:10.1006/jcsc.1998.0238.
144. Mason-Gamer RJ, Weil CF, Kellogg EA (1998) Granule-bound starch synthase: structure, function, and phylogenetic utility. *Mol Biol Evol* 15: 1658–1673. doi:10.1093/oxfordjournals.molbev.a025893. PubMed: 9866201.
145. Clark JR, Robertson M, Ainsworth CC (1991) Nucleotide sequence of a wheat (*Triticum aestivum* L.) cDNA clone encoding the waxy protein. *Plant Mol Biol* 16: 1099–1101. doi:10.1007/BF00016086. PubMed: 1863765.
146. Klosgen RB, Gierl A, Schwarz-Sommer ZS, Saedler H (1986) Molecular analysis of the waxy locus of *Zea mays*. *Mol Gen Genet* 203: 237–244. doi:10.1007/BF00333960.
147. Rohde W, Becker D, Salamini F (1988) Structural analysis of the waxy locus from *Hordeum vulgare*. *Nucleic Acids Res* 16: 7185–7186. doi:10.1093/nar/16.14.7185. PubMed: 2970062.
148. Shure M, Wessler S, Fedoroff N (1983) Molecular identification and isolation of the waxy locus in maize. *Cell* 35: 225–233. doi:10.1016/0092-8674(83)90225-8. PubMed: 6313224.
149. Wang ZY, Zheng FQ, Gao JP, Snustad DP, Li MG et al. (1995) The amylose content in rice endosperm is related to the post-transcriptional regulation of the waxy gene. *Plant J* 7: 613–622. doi:10.1046/j.1365-3113.1995.7040613.x. PubMed: 7742858.
150. Hunt HV, Moots HM, Graybosch RA, Jones H, Parker M et al. (2013) Waxy phenotype evolution in the allotetraploid cereal broomcorn millet: mutations at the GBSSI locus in their functional and phylogenetic context. *Mol Biol Evol* 30: 109–122. doi:10.1093/molbev/mss209. PubMed: 22936718.
151. Goh WL, Chandran S, Franklin DC, Isagi Y, Koshy KC et al. (2013) Multi-gene region phylogenetic analyses suggest reticulate evolution and a clade of Australian origin among paleotropical woody bamboos (Poaceae: Bambusoideae: Bambuseae). *Plant Syst Evol* 299: 239–257. doi:10.1007/s00606-012-0718-1.
152. Zhang Y-X, Zeng C-X, Li D-Z (2012) Complex evolution in Arundinarieae (Poaceae: Bambusoideae): incongruence between plastid and nuclear gene phylogenies. *Mol Phylogenet Evol* 63: 777–797. doi:10.1016/j.ympev.2012.02.023. PubMed: 22415014.
153. Essi L, Longhi-Wagner HM, de Souza-Chies TT (2008) Phylogenetic analysis of the *Briza* complex (Poaceae). *Mol Phylogenet Evol* 47: 1018–1029. doi:10.1016/j.ympev.2008.03.007. PubMed: 18439843.
154. Yang H-Q, Peng S, Li D-Z (2007) Generic delimitations of *Schizostachyum* and its allies (Gramineae: Bambusoideae) inferred from GBSSI and *trnL-F* sequence phylogenies. *Taxon* 56: 45–54.
155. Fortune PM, Schierenbach KA, Ainouche AK, Jacquemin J, Wendel JF et al. (2007) Evolutionary dynamics of *Waxy* and the origin of hexaploid *Spartina* species. *Mol Phylogenet Evol* 42: 1040–1055.
156. Bremer K, Wanntorp H-E (1979) Hierarchy and reticulation in systematics. *Syst Zool* 28: 624–627. doi:10.2307/2412572.
157. Nelson G, Platnick N (1980) Multiple branching in cladograms: two interpretations. *Syst Zool* 29: 86–91. doi:10.2307/2412630.
158. McDade L (1992) Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis. *Evolution* 46: 1329–1346. doi:10.2307/2409940.
159. Hedberg O (1995) Cladistics in taxonomic botany – master or servant? *Taxon* 44: 3–22. doi:10.2307/1222672.
160. Sosef MSM (1997) Hierarchical models, reticulate evolution, and the inevitability of paraphyletic supraspecific taxa. *Taxon* 46: 75–85. doi:10.2307/1224293.

161. Mindell DP (1991) Similarity and congruence as criteria for molecular homology. *Mol Biol Evol* 8: 897–900.
162. Mindell DP (1992) Phylogenetic consequences of symbioses: Eukarya and Eubacteria are not monophyletic taxa. *Biosystems* 27: 53–62. doi: 10.1016/0303-2647(92)90046-2. PubMed: 1391691.
163. Stebbins GL (1956) Taxonomy and the evolution of genera, with special reference to the family Gramineae. *Evolution* 10: 235–245. doi: 10.2307/2406009.
164. Kellogg EA (1989) Comments on genomic genera in the Triticeae (Poaceae). *Am J Bot* 76: 796–805. doi:10.2307/2444536.
165. Barkworth ME (2000) Changing perceptions of the Triticeae. In: SWL Jacobs, J Everett. *Grasses: Systematics and Evolution*. Melbourne: CSIRO. pp. 110–120.