Reticulate Evolutionary History of a Complex Group of Grasses: Phylogeny of Elymus StStHH Allotetraploids Based on Three Nuclear Genes

Roberta J. Mason-Gamer,* Melissa M. Burns†a, Marianna Naumb

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

Abstract

Background: Elymus (Poaceae) is a large genus of polyploid species in the wheat tribe Triticeae. It is polyphyletic, exhibiting many distinct allopolyploid genome combinations, and its history might be further complicated by introgression and lineage sorting. We focus on a subset of Elymus species with a tetraploid genome complement derived from Pseudoroegneria (genome St) and Hordeum (H). We confirm the species’ allopolyploidy, identify possible genome donors, and pinpoint instances of apparent introgression or incomplete lineage sorting.

Methodology/Principal Findings: We sequenced portions of three unlinked nuclear genes—phosphoenolpyruvate carboxylase, β-amylase, and granule-bound starch synthase I—from 27 individuals, representing 14 Eurasian and North American StStHH Elymus species. Elymus sequences were combined with existing data from monogenic representatives of the tribe, and gene trees were estimated separately for each data set using maximum likelihood. Trees were examined for evidence of allopolyploidy and additional reticulate patterns. All trees confirm the StStHH genome configuration of the Elymus species. They suggest that the StStHH group originated in North America, and do not support separate North American and European origins. Our results point to North American Pseudoroegneria and Hordeum species as potential genome donors to Elymus. Diploid P. spicata is a prospective St-genome donor, though conflict among trees involving P. spicata and the Eurasian P. strigosa suggests either introgression of GBSSI sequences from P. strigosa into North American Elymus and Pseudoroegneria, or incomplete lineage sorting of ancestral GBSSI polymorphism. Diploid H. californicum and/or allotetraploid H. jubatum are possible H-genome donors; direct involvement of an allotetraploid Hordeum species would simultaneously introduce two distinct H genomes to Elymus, consistent with some of the relationships among H-genome sequences in Hordeum and Elymus.

Conclusions/Significance: Comparisons among molecular phylogenetic trees confirm allopolyploidy, identify potential genome donors, and highlight cases of apparent introgression or incomplete lineage sorting. The complicated history of this group emphasizes an inherent problem with interpreting conflicts among bifurcating trees—identifying introgression and determining its direction depend on which tree is chosen as a starting point of comparison. In spite of difficulties with interpretation, differences among gene trees allow us to identify reticulate species and develop hypotheses about underlying evolutionary processes.

Introduction

Untangling reticulate relationships among species presents an interesting challenge to systematists, and an opportunity to uncover previously undetected evolutionary processes. Comparisons among gene trees can clarify historical relationships among species, and the examination of topological conflicts among trees can reveal complicating factors such as retention of ancestral genetic polymorphism, past or ongoing gene exchange, allopolyploidy, or a combination of these. Distinguishing among potential causes of phylogenetic conflict is often difficult, but careful comparisons among trees can help pinpoint the species involved, and allow specific hypotheses to be formulated. In the present study, we focus on species that are explicitly reticulate in that they are all allotetraploids, and potentially secondarily reticulate if they have arisen through multiple independent origins or undergone hybridization at the tetraploid level. We assess the role of reticulation at both levels in the genus Elymus L. of the wheat tribe Triticeae (Poaceae), using phylogenetic analyses of three unlinked, low-copy genetic markers.
Many inferences of reticulate evolution have been based on comparisons among gene trees in plants, comparisons between chloroplast DNA (cpDNA) and nuclear ribosomal internal transcribed spacer (ITS) phylogenies are especially widely used (e.g., [1–7]). The reasons are partly historical— for technical reasons, these high-copy markers were among the first to be widely used for plant phylogenetic studies. They remain in frequent use in studies of reticulation (e.g., [8–12]), where they offer both methodological advantages and a kingdom-wide backdrop of published sequences within which new data can be interpreted. However, both cpDNA and ITS sequence data sets have some disadvantages. The chloroplast genome is maternally inherited in most angiosperms, and its ability to identify a maternal donor can be an advantage. However, its inability to provide information about other genetic donors is often a major limitation. The biparentally-inherited ITS does have the potential to reveal multiple genome donors, but its arrangement in long repetitive arrays promotes the confounding effects of concerted evolution, both within arrays [13–15] and among them [16,17]. Thus, ITS copies can potentially convert toward one or the other parent, and the resulting sequence homogeneity can obscure a history of contributions from multiple distinct donors.

Low-copy nuclear genes, like ITS, reveal multiple genome donors, and they are less subject to gene conversion. However, they do have some practical disadvantages. They can be more difficult to amplify because of their low copy number, and because online databases often contain fewer comparable sequences from which amplification primers can be designed. The smaller sequence database also narrows the phylogenetic context within which new data sets can be analyzed, and makes it more difficult to assemble the crucial copy-number information that would prevent misinterpretation of unsuspected variation among paralogs. In spite of the difficulties, a variety of single- and low-copy nuclear genes have been successfully used in many studies of reticulate relationships in plants over the last decade (e.g., [18–32]). Sequence data from some low-copy genes are now becoming plentiful across a broad range of angiosperms.

This study presents three low-copy nuclear gene trees from a group of tetraploid species in the wheat tribe, Triticeae. The wheat tribe is especially well known for its economically important members, including wheat, barley, and rye. The tribe’s economic importance has driven an interest in its evolutionary history seemingly disproportionate to its size (about 300 species), yet a singular, straightforward phylogenetic estimate for the tribe seemingly disproportionate to its size (about 300 species), yet a singular, straightforward phylogenetic estimate for the tribe seems established before the start of the study, rather than a hypothesis in need of additional tests. Our main goals, therefore, were to (1) determine whether the North American vs. Eurasian StStHH species arose from separate polyploidization events; (2) identify possible progenitor species from within Hordeum and Pseudoroegneria; and (3) find out whether reticulate patterns in StStHH Elymus extend beyond those clearly attributable to allopolyploidy alone.

**Materials and Methods**

**Samples**

The analyses include 27 individuals representing 14 StStHH tetraploid *Elymus* species: 11 individuals representing five Eurasian species (Table 1), and 16 individuals representing nine North American species (Table 2). Most of the accessions were obtained from the USDA and have associated chromosome counts (2n = 28), confirming their tetraploid nature. Nearly all of the sequences from the Eurasian *Elymus* samples were newly generated for this study; the few exceptions [55] are noted in Table 1. Many of the North American *Elymus* pepC and GBSSI sequences were drawn from earlier studies [29,56], while the North American β-amylase sequences are new. The numbers of intraspecific samples of North American species differ among the three gene trees, but this does not affect our main conclusions.

Three single- or low-copy nuclear genes were amplified from each *Elymus* individual, and 8–24 clones per gene per individual were checked; the goal was to obtain copies from the St and H genomes from each. For each gene, the cloned sequences from *Elymus* were analyzed with previously-published sequences from a reasonably broad sample of the tribe’s known genomic diversity, including representatives of 15 monogenic genera (Table 3). These include the donor of the St genome (*Pseudoroegneria*), and of the genomes known to co-occur with St in allopolyploid nuclei: *Hordeum* (H), *Agropyron* (P), *Australopyrum* (W), *Psathyrostachys* (Ns; this genome is represented by tetraploid *Leymus* Hochst. in the pepC data set), and *Thinopyrum* (J and/or E). Additional monogenic genera include *Aegilops* L., *Crypsis* Jaub. & Spach (except for pepC), *Dasypyrum* (Coss. & Durieu) T.Durand, *Eremopyrum* (Ledeb.) Jaub. & Spach, *Henardia* C.E.Hubb. (except for pepC), *Heteranthelium* Hochst. ex Jaub. & Spach, *Peridictyon* Seberg, Fred., & Baden, *Secale* L., *Taeniatherum* Nevski, and *Triticeum*. The sample includes most of the monogenic genera accepted in genome-based classifications of the Triticeae (e.g., [35,36]). All three gene trees were rooted with a representative of *Bromus* L.; Bromeeae and Triticeae have repeatedly been shown to form a single clade, with Bromeeae as either sister or paraphyletic to a monophyletic Triticeae [57,58].

Nearly all of the sequences from the monogenic species were previously published in various sources [21,29,36,39–64], with a few exceptions as noted in Table 3. Information about the data and taxa can be found therein, but the primary discussions about
the characteristics of each marker and data set are: pepC [56]; β-amylase [61]; and GBSSI [29,63]. Based on studies of grass genomes in crop species, the three nuclear markers appear to be on three different chromosomes (more below). This assumption is tentative because it is based on a small number of grass species, but for this study the three genes are assumed to be unlinked, and to represent independent phylogenetic estimates.

Molecular methods and alignment

Similar molecular methods were followed for each of the three nuclear gene fragments (detailed protocols are found in the works cited above for each marker). For each Elymus individual, three PCR replicates were run per individual, in order to counter the potential effects of PCR drift [65]. PCR products from replicate reactions were combined and cleaned on columns (Qiagen). Cleared products were cloned into pGEM-T Easy vectors (Promega), and transformed into E. coli JM109 competent cells (Promega) following the manufacturer’s protocol, except that all reactions were halved. Cloned fragments were amplified directly from white colonies using the same primers as were used for the original PCR, in 30–40 μl reactions with 0.5 units Taq polymerase (Sigma), a 1× concentration of the included Taq buffer, 45–60 mmol MgCl₂, 6–8 nmol of each nucleotide, and 30–40 pmol of each primer. Amplified fragments were cleaned for sequencing using 1 unit shrimp alkaline phosphatase (USB) and 5 units exonuclease I (USB). Sequencing reaction included 1–3 μl of cleaned product, 2 μl BigDye Terminators v.3.1 (Applied Biosystems), and 3.2 pmol primer in a 10 μl volume. Reactions were run on an ABI Prism 377 (our lab) or ABI 3730 DNA Analyzer (Pritzker Lab, Field Museum of Natural History). Between 8 and 24 cloned PCR products from each individual were screened with a single sequencing primer, yielding a single-stranded partial sequence of about 600 basepairs. We examined these preliminary sequences to identify the two homoeologous sequence types (St and H) that we expected to find within each tetraploid individual. Representative clones of each were fully sequenced on both strands and added to the data set. If either homoeologous copy was missing from an initial sample of 8–12 clones, the corresponding gene from that individual was re-amplified and cloned, and 12 additional sequences were screened. We also included distinct, same-genome alleles from within individuals when they were encountered, although this was not our main goal. Based on the 600-basepair preliminary sequences, same-genome sequences that differed by more than three basepair substitutions were fully sequenced and added to the data set. The

### Table 1. Eurasian StStHH tetraploid Elymus species.

| Accession¹ | Country of Origin | #² | Genome | pepC  | b-amylase | GBSSI⁶ |
|------------|------------------|----|--------|-------|-----------|--------|
| Elymus brachyarratus PI 499411 | China | 1 | St     | b²-HM035290 | b-HM035223 | a-HM035268 |
|           |                  |    | H      | e-HM035291 | a-HM035224 | b-HM035269 |
| Elymus caninus PI 314205 Uzbekistan | PI 499413 | China | 4 | St     | d-HM035297 | d-HM035229 | a-HM035272 |
|           |                  |    | H      | f-HM035298 | b-HM035230 | b-HM035273 |
| PI 531571 Poland | 5 | St     | e-HM035299 | d-HM035231 | a-HM035274 |
| Elymus dentatus PI 628702 Russia | 1 | St     | d-HM035301 | a-HM035233 | a-DQ159328 |
|           |                  |    | H      | c-HM035302 | d-HM035234 | b-DQ159329 |
| PI 531599 Pakistan | 2 | St     | a-HM035304 | i-HM035235 | a-HM035276 |
|           |                  |    | H      | b-HM035305 | a-HM035236 | b-HM035277 |
| Elymus mutabilis PI 628704 Russia | 1 | St     | NR⁴ | a-HM035237 | c-DQ159331 |
|           |                  |    | H      | a-HM035306 | c-HM035238 | a-DQ159330 |
| PI 499449 China | 2 | St     | b-HM035308 | h-HM035239 | a-HM035278 |
|           |                  |    | H      | a-HM035309 | c-HM035240 | b-HM035279 |
| Elymus sibiricus PI 628699 Russia | 1 | St     | b-HM035311 | d-HM035231 | a-HM035280 |
|           |                  |    | H      | a-HM035312 | b-HM035241 | a-HM035281 |
| PI 499461 China | 3 | St     | d-HM035313 | g-HM035242 | a-HM035282 |
|           |                  |    | H      | a-HM035314 | b-HM035243 | b-HM035283 |

¹Plant introduction (PI) numbers were assigned by the USDA. Voucher specimens are at ID.
²Numbers (#) distinguish individuals within species on Figures 1–3, S1.
³Letters identify cloned sequences from within each individual on Figures 1–3, S1.
⁴NR: not recovered for a gene/individual.
⁵DQ1593## accessions: [55]; all others are new.

doi:10.1371/journal.pone.0010989.t001
three-basepair threshold was arbitrary, but we reasoned that it was
large enough to reveal distinct alleles rather than Taq errors.

PCR amplification of intra-individual variants can yield
chimeric sequences [66–68]. A small number of recombinants
were identified by inspection of alignments prior to phylogenetic
analysis; such sequences are often visible because the
St and H variants have length differences in some of the introns. A few more
were identified following closer examination of sequences on long
branches in preliminary maximum parsimony analyses. Such
sequences were confirmed as recombinant by inspection or if, when they were divided at the presumed point of recombination and analyzed as separate sequences, one portion was phylogenetically St-like and the other was H-like. Chimeric sequences were interpreted as PCR artifacts and removed from the analyses.

The pepC gene is a member of a three-copy family in grasses
[69]; the sequences used here appear to be homologous to the

| Accession | # | Genome | pepC | β-amylase | GBSSI |
|-----------|---|--------|------|-----------|-------|
| Elymus canadensis PI 578675 2 St | b-AY553248 | a-HM035244 | c-HM035284 |
| Elymus canadensis PI 531568 4 St | a-AY553242 | d-HM035245 | a-HM035285 |
| Elymus elymoides PI 531606 1 St | b-AY553249 | a-HM035246 | a2-AO109991 |
| Elymus glaucus RJMG 130 4 St | a-AY553244 | c-HM035247 | d2-AO109965 |
| Elymus glaucus W6 10215 6 St | b-AY553250 | - | a-AO10973 |
| Elymus hystrix Barkworth 97–87 1 St | e-HM035316 | a-HM035248 | a-AO10967 |
| Elymus lanceolatus W6 14220 1 St | a-AY553245 | a-HM035249 | d-HM035286 |
| Elymus lanceolatus W6 14218 2 St | - | c-HM035250 | a-AO10993 |
| Elymus lanceolatus PI 531623 3 St | b-AY553252 | - | a-AO10973 |
| Elymus riparius RJMG 160 1 St | - | a-HM035254 | NR |
| Elymus trachycaulus PI 372500 1 St | e-HM035317 | a-HM035256 | b1-AO109086 |
| Elymus trachycaulus PI 452446 3 St | a-HM035318 | b-HM035257 | a1-AO109092 |
| Elymus virginicus RJMG 161 4 St | - | a-HM035259 | a-AO10974 |
| Elymus virginicus RJMG 168 9 St | a-HM035220 | e-HM035260 | b-a-AO10976 |
| Elymus wawawaiensis PI 285272 1 St | - | a-HM035261 | c-AO10996 |
| Elymus wawawaiensis PI 598812 3 St | b-AY553253 | c-HM035263 | a-AO10990 |
| Elymus wawawaiensis PI 598812 3 St | a-AY553247 | b-HM035264 | d-AO10978 |

1 Plant introduction (PI) and W6 numbers were assigned by the USDA; RJMG and Barkworth accessions are from the first author and M. Barkworth, respectively. The E. trachycaulus accessions were collected in Canada; all others are from the United States. Voucher specimens are at ID.

2 Numbers (###) distinguish individuals within species on Figures 1–3, S1.

3 Letters identify cloned sequences from within each individual on Figures 1–3, S1.

4 Dashes indicate individuals that were not sampled for the corresponding gene.

5 NR: not recovered for a gene/individual.

6,7 AY5532### accessions: [56]; AY0109### accessions: [29]; all others are new.

doi:10.1371/journal.pone.0010989.t002
| Species name | Accession | #   | pepC      | β-amylase | GBSSI   |
|--------------|-----------|-----|-----------|-----------|---------|
| *Aegilops* bicornis (Forskål) Jaub. & Spach. | Morrison s.n. | -   | -         | 6AYB1686 | -       |
| *Aegilops caudata* L. | G 758 | -   | 6AYB1687  | 6AF079262 |         |
| *Aegilops comosa* Sibth. & Smith | G 602 | 1AY533236 | 6AYB1690 | -         |
| *Aegilops speltoides* Tausch | Morrison s.n. | -   | -         | 6AF079267 |         |
| *Aegilops tauschii* Coss. | G 1297 | -   | 6AYB1691  | 6AF079270 |         |
| *Agropyron cristatum* (L.) Gaertn. | Pl 279802 | 1   | 6AYB1697  | 6AF079271 |         |
| *Agropyron cristatum* | Pl 281862 | 2   | -         | 1AY11002  |         |
| *Agropyron mongolicum* Keng | D 2774 | -   | 6AYB1693  | 1AY11004  |         |
| *Australopyrum retrofractum* (Vickery) Å. Löve | Pl 530013 | 1AY553238 | -         |         |         |
| *Australopyrum velutinum* (Nees) B.K. Simon | D 2873–2878 | 1   | 6AYB1698  | 6AF079274 |         |
| *Dasypyrum villosum* (L.) Candargy | PI 251478 | 1   | 6AYB1699  | -         |         |
| *Dasypyrum villosum* | Pl 470279 | 2   | -         | -         |         |
| *Dasypyrum villosum* | D 2990 | 3   | 1AY553240 | -         |         |
| *Eremopyrum banaevarius* (Spreng.) Nevski | H 5554 | -   | 6AYB1700  | 1AY11005  |         |
| *Eremopyrum distans* (C.Koch) Nevski | H 5552 | -   | 6AYB1701  | 1AY11006  |         |
| *Eremopyrum orientale* (L.) Jaub. & Spach | H 5555 | 1AY553254 | 6AYB1702  | 1AY11007  |         |
| *Herendia persica* (Boiss.) C.E. Hubb. | H 5556 | -   | 6AYB1703  | 6AF079276 |         |
| *Heteranthelium piliatum* (Banks & Sol.) Hochst. | PI 402352 | 1   | 6AYB1704  | 6AF079277 |         |
| *Hordeum bogdani* Wilensky | Pl 531762 3 | 1 | - | 2EU282293 | 2EU282255 | 2EU282317 |
| *Hordeum bogdani* | Pl 531760 3 | 2 | 2EU282293 | 2EU282255 | 2EU282317 | 2EU282293 |
| *Hordeum brevisubulatum* (Trin.) Link | PI 401387 | 1 | 6AYB1705  | 1AY10961  |         |
| *Hordeum californicum* Covas & Stebbins | MA-138-1-40 | 1 | 6AYB1706  | 1AY10962  |         |
| *Hordeum chilense* Roem. & Schult. | PI 531781 11 | 1 | 2EU282294 | 2EU282295 | 2EU282296 |
| *Hordeum jubatum* | Pl 531781 11 | 2 | 2EU282298 | 2EU282298 | 2EU282319 |
| *Hordeum marinus* | PI 304346 11 | 3 | 2EU282300 | 2EU282300 | 2EU282300 |
| *Leymus racemosus* ssp. sabulosus (M.Bieb.) Tzvel. | PI 531813 | 1 | 1AY553260 | 2EU282263 | -         |
| *Psathyrostachys fragilis* (Boiss.) Nevski | PI 343192 | -   | 6AYB1715  | 6AF079279 |         |
widely-expressed housekeeping copy. Based on the location of similar sequences in the rice genome (Genbank AP005781 and AP005802) and a comparative grass genome map [70], this gene copy is assumed to be on the Triticeae group 5 homoeologous chromosomes. The original Triticeae pepC data set [56] combined two fragments designated region 1 (approximately 1 kb; Genbank AY553236–AY553269) and region 2 (approximately 600 bp; Genbank AY548399–548432); the present data set includes just region 1 sequences. The 1100-bp PCR products obtained using primers 467F(1) and 1672R(2) [56] include partial exons 1 and 2, along with the intervening intron, which is approximately 1000 bp long. The intron exhibits considerable length variation, including insertion and excision of transposons [21]. Most length variation could be accommodated by manually adjusting the alignment (File S1). An ambiguous region of the alignment consisting mainly of short runs of C and/or T (positions 67–109), and two regions affected by transposon activity (690–771 and 1035–1119), were excluded from the analysis.

| Species name | Accession | # | pepC | β-amylose | GBSSI |
|--------------|-----------|---|------|-----------|-------|
| *Psathyrostachys juncea* (Fisch.) Nevski | Pi 206684 | - | 4AY821716 | 4AF079280 |
| *Pseudoroegneria libanotica* (Hack.) D.R.Dewey | Pi 228391 | 1 | 2EU282304 | 2EU282264 | 2EU282324 |
| *Pseudoroegneria libanotica* | Pi 228392 | 3 | 2EU282305 | 2EU282265 | 2EU282325 |
| *Pseudoroegneria spicata* (Pursh) A. Löve subsp. spicata | Pi 232117 | 1 | 4AY821717 | 4AF079281 |
| *Pseudoroegneria spicata* subsp. *inermis* (Scribn. & J.G.Smith) A. Löve | Pi 236681 | 2 | 4AY821718 | 4AY010998 |
| *Pseudoroegneria spicata* subsp. *spicata* | Pi 610986 | 3 | 1AY553263 | - | 4AY010999 |
| *Pseudoroegneria spicata* subsp. *spicata* | D 2844 | 4 | 1AY553264 | 4AY821719 | 4AY011000 |
| *Pseudoroegneria stipifolia* (Czern. ex Nevski) A. Löve | Pi 313960 | 2 | 2EU282306 | 2EU282266 | - |
| *Pseudoroegneria stipifolia* | Pi 531751 | 3 | 2EU282307 | 4AY821721 | - |
| *Pseudoroegneria stipifolia* (M.Bieb.) A. Löve | Pi 499637 | 1 | 2EU282309 | - | 2EU282323 |
| *Pseudoroegneria stipifolia* ssp. aegilopoides (Drobow) A. Löve | Pi 531755 | 2 | 2EU282311 | 2EU282267 | 4AY360823 |
| *Pseudoroegneria tauri* (Boiss. & Balansa) A. Löve | Pi 380652 | 1 | 2EU282312 | 2EU282268 | 2EU282326 |
| *Pseudoroegneria tauri* | Pi 401319 | 2 | 2EU282313 | - | 2EU282327 |
| *Pseudoroegneria tauri* | Pi 380644 | 3 | 2EU282314 | - | - |
| *Secale cereale* L. | Kellogg s.n. | 1 | 1AY553266 | 4AY821723 | 4AY011009 |
| *Secale montanum* (C.Presl.) C.Presl. | T 36554 | - | 4AY821725 | - |
| *Secale montanum* | Pi 440654 | - | - | 4AF079282 |
| *Secale strictum* subsp. anatolicum (Boiss.) K.Hammer | Pi 206992 | 1 | 1AY553265 | 4AY821722 | 4AY011008 |
| *Taeniatherum caput-medusae* (L.) Nevski | Pi 208075 | 1 | - | 4AY821726 | 4AY011010 |
| *Taeniatherum caput-medusae* | RJMG 189 | 2 | 1AY553268 | 4AY821727 | 4AY360847 |
| *Taeniatherum caput-medusae* | RJMG 113 | 2 | 1AY553269 | - | 4AY360848 |
| *Thinopyrum bessarabicum* (Savul. & Rayss) A. Löve | Pi 314697 | 3 | - | 4AY821728 | - |
| *Triticum aestivum* L. | 3AJ007705 | - | - | - |
| *Triticum aegilopicum* Boiss. | Morrison s.n. | - | 4AY821732 | 4AF079285 |
| *Triticum monococcum* L. | Pi 221413 | - | 4AY821733 | - |
| *Triticum urartu* Tumanian | Morrison s.n. | - | 4GQ847677 | 4AF079287 |
| *Bromus tectorum* L. | Kellogg s.n. | 1 | 1AY553239 | 4AY821734 | 4AY362757 |

1–9 From [56,21,59,61,62,63,29,64,60], respectively.
10 New to this study.
11 Indicated accessions are deposited at ID; others are at GH.
doi:10.1371/journal.pone.0010989.t003

Table 3. Cont.
The β-amylase genes form a small family in the Triticeae, with several copies expressed in the endosperm and one that is ubiquitously expressed [71]. Based on sequence similarity, the sequences used here appear to represent the ubiquitously-expressed copy; this copy has been mapped to the Triticeae group 2 homoeologous chromosomes [72]. The 1400-bp β-amylase PCR products were obtained using primers 2a-for and S5bac [61], and include partial exons 2 and 5, complete exons 3 and 4, and introns 2–4, which are about 250, 100, and 400 bp in length, respectively. The β-amylase alignment (File S2) was generally straightforward; most length differences were easy to interpret. One ambiguous simple sequence region (positions 553–570) and two regions corresponding to Stowaway-like transposon activity in some sequences (positions 635–765 and 1475–1641) [73] were excluded from the analyses. The GBSSI PCR products were obtained using the F-for and M-bac primers [63], which amplify an approximately 1300-bp fragment that includes partial exons 9 and 14, exons 10–13, and introns 9–13, which are about 100 bp each. The putatively single-copy GBSSI gene maps to the Triticeae group 7 homoeologous chromosomes [70,74], or to a portion of chromosome 4 translocated from, and thus homoeologous to, the group 7 chromosomes [70,75]. The GBSSI alignment (File S3) is generally straightforward in spite of numerous small insertions and deletions in the introns. Three ambiguously-aligned regions (positions 910–1020, 1138–1231, and 1383–1515) were excluded from the phylogenetic analyses.

Phylogenetic analyses

Prior to phylogenetic analyses, 16 nested models of sequence evolution [76–78] were examined for each data set using preliminary maximum parsimony trees, and the resulting maximum-likelihood (ML) scores were compared using a likelihood ratio test [78–81]. For each data set, the general time-reversible (GTR) substitution model [82,83] led to a large and significant increase in score compared to the Jukes-Cantor [84], Kimura two-parameter [85], and Hasegawa-Kishino-Yano [86] models, as did the addition of a gamma (Γ) distribution with shape parameter 2 to model among-site rate variation [87]. Adding an invariable sites (I) parameter [86] to the GTR+Γ model led to a non-significant increase in the pepC and β-amylase scores, and a significant increase in the GBSSI score. Therefore, the GTR+Γ model was used for the ML analysis of the pepC and β-amylase data, and the GTR+I+Γ model was used for the analyses of the GBSSI data sets.

All ML analyses were run using the Mac OS X UNIX version of GARLI v. 0.95 [88]. Following the recommendations of the author, runs were set for an unlimited number of generations, and automatic termination following 10,000 generations without a significant [lnI. increase of 0.01] topology change. For each data set, thirty analyses were run with random starting tree topologies, and the tree with the best score was used to display the gene tree. Branch support (BS) for each tree was estimated based on 100 ML bootstrap replicates in GARLI with searches as above, except that the stopping criterion was lowered to 5,000 generations without a significant topology change. Bootstrap values ≥50% were recorded on the best ML trees.

Results

General

On all three gene trees (Figures 1–3) the Elymus sequences fall into two main groups, with Pseudoroegneria and with Hordeum. The two distinct Elymus groups are interpreted as St and H sequences, respectively, derived from Pseudoroegneria and Hordeum progenitors. Nearly all of the Elymus individuals yield sequences in both of the main clades, with the following exceptions: for pepC, only an H sequence was recovered from E. virginicus 4 and E. mutabilis 1; for β-amylase, only an H sequence was recovered from E. glaucus 6, E. hystrix 1, E. virginicus 4, and E. sibiricus 1; and for GBSSI, only an H sequence was recovered from E. riparius 1. No individuals lack a sequence type in all three data sets, and only Elymus virginicus 4 is missing a sequence type for two of the three genes. There are cases where individual plants show within-genome (presumably allelic) variation, and these cases are used where possible to shed light on the evolutionary history of the StStHH Elymus group. However, the study was designed primarily to capture intergenomic variation within individuals (i.e., St vs. H) rather than allelic variation, so patterns of intra-individual variation within genomes are probably more widespread than the data show.

There are several cases of apparent β-amylase homolog silencing in Elymus, inferred from exon sequences. Of the Eurasian species, two of the four E. caninus individuals have a stop codon in exon 2 of their St-genome copies (clones 2d and 4d). Silent copies are more widespread among the North American individuals, and all involve the H-genome copy. The E. hystrix 1a, E. riparius 1b, and E. virginicus 9e clones share a 2-basepair deletion in exon 2, and E. hystrix 1a has a second, single-basepair deletion in exon 4; the E. eymoides 1c and E. glaucus 6a clones each have a unique, single-basepair deletion in exon 2; and E. varvovisaiis 3b has a stop codon in exon 4. Note that for two individuals (E. glaucus 6 and E. hystrix 1), no functional β-amylase copy was recovered; in both cases, the St copy was not present among the sequenced clones, and the H copy includes a frame-shifting deletion.

The pepC phylogeny

The St-genome sequences of Pseudoroegneria and Elymus form a well-supported clade (100% BS; Figure 1). The Elymus St sequences show little diversity overall, and are most closely related to the only native North American Pseudoroegneria species, P. spicata (85% BS), from which they show very little divergence. The only phylogenetic structure within the clade groups the two E. sibiricus sequences with E. brechyiannis (70% BS), and one of the two E. dentatus sequences with E. mutabilis (84% BS). There is no phylogenetic distinction between the North American and Eurasian species.

The remaining Elymus pepC sequences represent the H-genome, and form a clade with part of Hordeum (Figure 1). The Elymus H-genome sequences, like the St sequences, show little diversity. They form a clade (78% BS) with sequences from H. californicum, a native North American diploid species, and with H. jubatum, a North American allotetraploid whose ITS sequences were derived from H. californicum and H. roheintzii [89]. This clade includes three subgroups: most of Elymus along with H. californicum and one of the H. jubatum genomes (77% BS); the four remaining Elymus sequences (74% BS); and the second H. jubatum genome (95% BS). The four individuals in the smaller Elymus clade (E. mutabilis 1 and 2, E. dentatus 1, and E. caninus 2) are also represented in the larger H clade; thus, the small clade appears to reveal allelic variation in the H genome. The three species in this clade are Eurasian natives, so the allele may be restricted to Europe, though more intensive sampling would be required to support this.

The β-amylase phylogeny

Within the St-genome clade on the β-amylase tree (Figure 2; 99% BS), the Elymus sequences are again most closely related to P. spicata, though the support for this relationship is only moderate (72% BS). Compared to the pepC tree, there is more phylogenetic
structure among the \(\beta\)-amylose \(St\) sequences of \textit{Elymus} and \textit{P. spicata}, including weak (64\% BS) support for a Eurasian species clade within a paraphyletic North American \textit{Pseudoroegneria/Elymus} assemblage. Intraspecific sampling is limited, but within the Eurasian clade, the \(St\) sequences representing \textit{E. dentatus} 1 and 2 form a monophyletic group (62\% BS), while those from \textit{E. caninus} and from \textit{E. mutabilis} do not. The \(St\) sequences representing the North American species \textit{E. wawawaiensis} are non-monophyletic, and those from \textit{E. lanceolatus} 1 and 2 are unresolved; the remaining species only have a single representative on the tree.

In contrast to the pepC tree, the \(H\) sequences from \textit{Elymus} do not group with the diploid \textit{H. californicum} on the \(\beta\)-amylose tree, but only with one of the \(H. jubatum\) genomes (100\% BS). The \(H. californicum\) sequence forms a clade with the other \(H. jubatum\) genome (100\% BS) within a large, multi-species \textit{Hordeum} clade (92\% BS) sister to (97\% BS) the \textit{Elymus}/\textit{H. jubatum} clade. There is much less resolution among the \(Elymus\) \(St\) sequences than among the \(St\) sequences. The few relationships with >50\% BS support are within-continent groups, but there is no suggestion of a Eurasian species clade, as there is in the \(St\)-sequence group.

The GBSSI phylogeny

The structure of the \(St\)-genome clade on the GBSSI tree differs from that on the other two trees. The group includes a paraphyletic “core” assemblage on short branches, in which the \textit{Elymus} sequences are similar to those from five \textit{P. spicata} individuals (1, 2, 3, 4, and 6). Within this assemblage, the Eurasian \textit{Elymus} sequences form a moderately supported clade (76\% BS) derived from within a paraphyletic group that includes \textit{P. spicata} and six sequences from four North American \textit{Elymus} species. This pattern by itself is similar to that in the \(\beta\)-amylose \(St\) clade, and is suggestive of a single origin of Eurasian species from within a paraphyletic group of North American \textit{Elymus} and \textit{Pseudoroegneria} species. The GBSSI \(St\) clade, however, is unique in having a separate subclade (87\% BS) on a relatively long branch nested within the core group. The subclade includes the second of two sequences from \textit{P. spicata} 6, the remaining ten North American \textit{Elymus} sequences, and, surprisingly, two Chinese accessions of \textit{P. sinigoua}. Like \textit{P. spicata} 6, three \textit{Elymus} individuals have gene copies in both the paraphyletic assemblage of short branches, and in the long-branched subclade: \textit{E. lanceolatus} 1 (clones a and d) and 2 (d and c), and \textit{E. wawawaiensis} 3 (b and a). At the species level, \textit{E. virginicus} is also represented in both groups, with individual 4 separated from individual 9. Another unique feature of the GBSSI tree is the placement of \textit{P. libanotica} and \textit{P. tauri} far outside of the main \(St\)-sequence clade, though their position within the tree is not convincingly resolved. The placement of these sequences as a possible result of diploid-level introgression, and their contribution to some of the species in a different group of \textit{Elymus} tetraploids (genomes \textit{StStYY}) is discussed elsewhere [62]. Based on the present sample, however, these species play no role in the evolution of the \textit{StStHH} \textit{Elymus} tetraploids.

The GBSSI \(H\)-genome clade, like the \(St\)-genome clade, is more complex than its counterparts on the pepC and \(\beta\)-amylose trees. All but one of the North American \textit{Elymus} sequences form a well-supported clade (100\% BS), along with \textit{H. californicum}, one of the tetraploid \textit{H. jubatum} genomes, and two Eurasian species (\textit{E. dentatus} 1 and 2, and \textit{E. mutabilis} 2). The remaining, mostly-Eurasian \textit{Elymus} sequences form a very weak (28\% BS) group with much greater sequence diversity. This assemblage includes a small clade of sequences from \textit{E. sibiricus} 1 and 3 and \textit{E. mutabilis} 1 (100\% BS), and a larger clade (93\% BS) with the second \textit{H. jubatum} genome, two Eurasian \textit{Elymus} species (\textit{E. brachyanistatus} and \textit{E. caninus}), and one accession of the North American species \textit{E. lanceolatus}. The high diversity among the GBSSI \(H\)-genome sequences, compared to those on the pepC and \(\beta\)-amylose trees, was unexpected enough to raise suspicion about previously undetected alignment artifacts in the introns. There is considerable length variation in the GBSSI introns, so we ruled out the possible effects of non-orthologous intron alignment by running a ML analysis of just the \(H\)-genome clade using the exons only, for which the alignment is unambiguous. The resulting tree (Figure S1) has a similar (though less well-resolved) topology, including the same clade of very similar, mostly North American sequences, with the remaining mostly-Eurasian sequences forming a weak group of long branches outside of the main \textit{Elymus} clade.

Discussion

The \textit{StStHH} genome configuration

Our initial assumption that the \textit{Elymus} species included here are \textit{StStHH}-genome tetraploids was based on the results of numerous studies of meiotic chromosome pairing (e.g., [54] and references therein). As expected, all three gene trees unequivocally support the \textit{Pseudoroegneria} + \textit{Hordeum} origin of the sampled \textit{Elymus} species, in agreement with an earlier analysis of RPB2 gene sequences [90]. Nearly all of the \textit{Elymus} individuals have two distinct copies of all three genes, in clades with \textit{Pseudoroegneria} and \textit{Hordeum}. No individuals are missing either copy of all three genes; thus, the occasional missing copies from a few individuals might represent sampling artifacts, copy loss, or unique changes in primer sites, but do not suggest that either genome is absent altogether.

The pepC and \(\beta\)-amylose trees

The pepC and \(\beta\)-amylose trees are in general agreement with regard to the relationships among the \textit{Elymus} species. Both show the North American and Eurasian species to be very closely related, and neither support independent origins for the two geographic groups. The \(St\) and \(H\)-genome clades on both trees support a North American origin for the \textit{StStHH} \textit{Elymus} tetraploids. Of the \textit{Pseudoroegneria} and \textit{Hordeum} species included in the analysis, those most closely related to \textit{Elymus} are North American species. Furthermore, within the \(St\)-sequence clade on the \(\beta\)-amylose tree, the Eurasian \textit{Elymus} species form a clade (albeit weakly supported) within a broader paraphyletic assemblage of North American \textit{Elymus} and \textit{Pseudoroegneria} sequences. Evidence for a single, North American origin of the \textit{StStHH} tetraploids is at odds with the earlier suggestion that the North American and Eurasian species arose separately, based on limited karyotype [91] and isoenzyme data [92,93]. More recently, separate origins were moderately well-supported in a phylogenetic analysis of a nuclear RNA polymerase II (RPB2) gene [90]; \(H\)-genome RPB2 sequences separated \textit{Elymus} into largely American and Eurasian subclades with, respectively, American and Eurasian \textit{Hordeum}
β-amylase

GTR + Γ

Tetraploid Elymus Phylogeny

[Diagram showing a phylogenetic tree with species names and branch lengths.]
species. The RPB2 St-genome sequences did not reveal a geographic pattern, and were ambiguous with regard to whether the closest Pseudoroegneria species was North American or Eurasian, placing one accession each of *P. spicata* and *P. stipifolia* within *Elymus*, and a second accession of *P. spicata* outside of the *Elymus* clade on a very long branch. (While all of our trees implicate *P. spicata* as a potential donor to the *StStHH* tetraploids, none point to *P. stipifolia*).

The pepC and β-amylase trees do differ with respect to the relationships between *Elymus* and * Hordeum*; specifically, the possible roles of diploid *H. californicum* and allotetraploid *H. jubatum* in the origin of *StStHH* *Elymus*. On the pepC tree, the *H*-genome sequences of *Elymus* are grouped with, and very similar to, the North American diploid *H. californicum*, and with both genomes of the allotetraploid *H. jubatum*. This is consistent with the results of an earlier study of repetitive DNA sequences [94], and one based on starch synthase data [29], both of which suggested *H. californicum* as a possible *H*-genome donor to *Elymus*. However, in contrast to the pepC tree, the β-amylase tree does not place *Elymus* with *H. californicum*, but instead with one of the *H. jubatum* genomes, while *H. californicum* is grouped with the other *H. jubatum* genome in a separate, multi-species *Hordeum* clade. Together, these trees and the differences between them suggest that a tetraploid similar to *H. jubatum* might have been involved in the history of *Elymus*, whether through past introgression among the *Elymus*, *H. californicum*, and *H. jubatum* lineages, or through a direct contribution from a tetraploid *H. jubatum*-like species to *Elymus*. Direct involvement of an *H. jubatum*-like ancestor would have led to the simultaneous introduction of both of its homoeologous *H* genomes, and in a successfully diploidized *StStHH* tetraploid, they would then behave as homologous alleles. Thus, depending on changes in allele frequency through time, *Elymus* might exhibit one or the other, or both, of the *H. jubatum*-like homoeologs.

The relationships among *H*-genome sequences could, instead, reflect introgression following tetraploid *Elymus* formation. It is impossible to trace a precise sequence of events, but we can envision scenarios consistent with the data. For example, if *H. californicum* was, in fact, the *H*-genome donor to *StStHH* *Elymus*, as suggested by the pepC tree, then *H. californicum* could be “misplaced” on the β-amylase tree, having acquired its β-amylase gene copy through introgression after the formation of *Elymus*. Alternatively, *Elymus’s* placement on the β-amylase tree, far from *H. californicum*, might indicate that it was *Elymus’s* β-amylase gene that was acquired through introgression; its close relationship to the second genome of *H. jubatum* indicates that species as a potential source. Additional samples of *H. californicum* and *H. jubatum* might support or refute our hypotheses, or suggest other conceivable scenarios, but in any case, it appears that *H. jubatum* was involved at some stage in the history of *StStHH* *Elymus*.

The GBSSI tree suggests further introgression or lineage sorting

If *Elymus* relationships are in some ways similar on the pepC and β-amylase trees, and relatively straightforward to interpret, the GBSSI results complicate the interpretation. In the *St* sequence clade, the paraphyletic “core” group of very similar sequences is, by itself, reminiscent of the pattern on the β-amylase tree, with the Eurasian species forming a moderately-supported clade within a paraphyletic group of North American *Elymus* and *Pseudoroegneria* species. This group, when considered alone, supports a North American origin of the *StStHH* tetraploid group from a *P. spicata*-like ancestor. However, the long-branched clade that arises from within the core group is unique to the GBSSI tree. It includes sequences from several North American *Elymus* individuals, some of which are also represented in the core group, and two Chinese accessions of a Eurasian *Pseudoroegneria* species, *P. strigosa*. The clade also includes one *P. spicata* sequence; the same individual (#6) has a second sequence in the core group with the rest of *P. spicata*. The dual placement of several individuals in both the core group and in the derived clade could reveal gene duplication. However, such a duplication event (at the base of the “St” clade; Figure 3) should also be evident in *P. strigosa* and in the Eurasian *Elymus* species, unless we postulate at least two subsequent, independent paralog losses. Thus, a more parsimonious explanation is that the relationships among GBSSI sequences from *P. spicata*, *P. strigosa*, and *Elymus* result from either past gene exchange or from the maintenance of a shared ancestral polymorphism. Introgression of a *P. strigosa*-like GBSSI allele into North America could explain the close relationship between *P. strigosa* and some of the sequences from *P. spicata* and North American *Elymus*, though the exact sequence of events is not clear from the present sample. The *P. strigosa* allele might have been introduced to North America through hybridization between *P. spicata* and *P. strigosa*, and then passed from *P. spicata* to North American *Elymus* through hybridization, or through formation of new *StStHH* tetraploids and subsequent hybridization among tetraploid lineages. The allele might have been introduced directly into *Elymus* through hybridization with a tetraploid (*StStSt*) accession of *P. strigosa*. Given the possible Eurasian origin of the allele, we could also postulate an introduction via Eurasian *Elymus* tetraploids, but so far, none of the sampled Eurasian *StStHH* *Elymus* species have alleles in this clade. Alternatively, the *P. spicata* GBSSI polymorphism, including the allele in the core clade relationship with the *P. strigosa* sequences, could reflect the maintenance of ancestral polymorphism as a result of incomplete lineage sorting. The subsequent introduction of both alleles from *P. spicata* into North American *Elymus* is consistent with the placement of *Elymus* alleles with both *P. strigosa* and *P. spicata* on the GBSSI tree. The Eurasian *Elymus* species lack the polymorphism; assuming this is not merely a sampling artifact, it appears that the *P. strigosa*-like allele was either never introduced into the Eurasian group, or that it was subsequently lost.

In the GBSSI *H*-clade, there is, again, a “core” group of very similar sequences (though monophyletic in this case) that includes most of the North American and a few Eurasian *Elymus* sequences with *H. californicum* and one genome of the tetraploid *H. jubatum*. The relationships among the *Elymus* sequences in the core clade once again suggest a North American origin, with a clade of Eurasian sequences nested within a paraphyletic North American *Elymus* and *Hordeum* group. However, the arrangement of *Elymus* sequences outside of the core clade bears no resemblance to either of the other trees, or to the *St*-genome clade on this tree. These species are primarily Eurasian, except for one of the two accessions of *E. lanceolatus*. The only *Hordeum* sequences loosely associated with these *Elymus* sequences represent the second genome from the tetraploid *H. jubatum*; this provides some additional support that...
Granule-bound Starch Synthase

GTR + I + Γ
H. jubatum was somehow involved in the history of the StStHHH Elymus species. This interpretation is not entirely satisfying, however, because the pattern is so unlike that on either of the other trees; these Elymus species are unexpectedly divergent from H. jubatum and from one another. Thus, there is no clear indication of a donor-recipient relationship among the sequences outside the core clade, whether from introgression or otherwise, so it is difficult to speculate on processes that might explain the topology of H-genome clade on the GBSSI tree. Consideration of the earlier RPB2 analysis [90] further complicates the interpretation. The Eurasian and American H-genome clades uncovered in that study were grouped with American and Eurasian representatives of Hordeum – H. stenostachys and H. bogdani, respectively – but these were the only two Hordeum species included in the RPB2 analysis, so a potential role of H. californicum and H. jubatum in the origin of Elymus was not assessed.

Summary and Conclusion

Our first goal was to test the hypothesis that the North American and Eurasian Elymus species arose independently; this was suggested by allozyme [92,93] and cytological [91] studies, and more explicitly supported by a recent molecular phylogenetic analysis of RPB2 [90]. However, our data support a single, North American origin. In the studies using cytological and allozyme data, issues of small sample size or the interpretation of complicated results could be invoked to explain away the contradiction, but the RPB2 results are more difficult to square with ours. We stand by the conclusion of a single origin because none of our three trees support separate origins. However, there are clearly other processes in play beyond the number of StStHH origins, and we have only four nuclear gene trees (including RPB2) in consideration, so additional trees could tip the interpretation in another direction. Our second goal, the identification of possible progenitors within Pseudoroegneria and Hordeum, yielded reasonably consistent results. Pseudoroegneria spicata is supported as the most likely St-genome donor among the species included on all of our trees (albeit with complications involving introgression or incomplete lineage sorting on the GBSSI tree), and as a possible progenitor on the RPB2 tree. Hordeum californicum and/or its tetraploid derivative H. jubatum are suggested as H-genome donors to Elymus on all of our trees; past interactions among these species remain to be clarified with more targeted sampling of these species and their closest relatives within Hordeum, and additional gene trees. Finally, our third goal was to identify reticulate patterns beyond allopolyploidy. Examination of the GBSSI tree relative to the others revealed a case of introgression or incomplete lineage sorting, revealed by the discovery of P. striigosa-like St-genome allele in North American Elymus and Pseudoroegneria, and a more confusing situation involving the GBSSI H-genome.

Our attempts to propose evolutionary scenarios to reconcile conflicting patterns among Elymus gene trees reveal a general problem with inferring reticulate events from multiple conflicting trees. The interpretation depends heavily on which tree is initially assumed to be closest to the “best” or “true” tree (with or without a clear explanation), after which the differences on the remaining trees are attributed to processes such as introgression or incomplete lineage sorting. In other words, the sequence of historical events leading to gene tree conflict is defined by which tree is selected as a reference tree, or the one with which other trees are viewed to be in conflict. (Perhaps the most familiar examples in plants involve conflicts between ITS and cpDNA trees, which are generally interpreted as cpDNA introgression. In that case, the ITS tree is being used as the reference tree, sometimes without explicit justification.) Furthermore, if there are numerous conflicts among trees, a different tree could potentially be selected as the reference tree for each group of conflicting taxa. If many gene trees are available, and one branching pattern is shared by a large majority of them, then it is probably reasonable to interpret the differences on the few conflicting trees with reference to the majority of trees. With just three trees presented here (or four, including RPB2 [90]), and without a clear majority pattern, the distinction between “real” and “conflicting” relationships is not always straightforward. On one hand, for example, the unexpected placement of some Pseudoroegneria and Elymus St sequences with P. striigosa on the GBSSI tree, in conflict with the pepC and β-amylase trees, seems like a fairly straightforward case in which the GBSSI gene was affected by introgression or maintenance of an ancestral polymorphism. On the other hand, an assumption that H. californicum is the diploid H-genome donor to StStHHH Elymus based on the pepC and GBSSI trees, and consequent interpretation of the β-amylase and RPB2 trees in terms of H-genome introgression, is more arbitrary. A proposed evolutionary scenario would be quite different if either the β-amylase or the RPB2 tree were assumed to represent the true tree with respect to H-genome sequence relationships. In spite of the difficulties with interpreting gene tree conflict as a specific sequence of evolutionary events, such patterns can certainly highlight the importance of such events, pinpoint the taxa involved, and yield hypotheses to be tested with targeted sampling within and among conflicting taxa, and with data from additional nuclear loci.

Supporting Information

File S1  Sequence data set - phosphoenolpyruvate carboxylase. Found at: doi:10.1371/journal.pone.0010989.s001 (0.15 MB TXT)

File S2  Sequence data set - beta-amylase. Found at: doi:10.1371/journal.pone.0010989.s002 (0.21 MB TXT)

File S3  Sequence data set - granule-bound starch synthase I. Found at: doi:10.1371/journal.pone.0010989.s003 (0.19 MB TXT)

Figure S1  Gene tree based on granule-bound starch synthase exon sequences. The best-scoring ML tree was selected from 30 GARLI analyses of GBSSI exons under a GTR+I+F model of sequence evolution. The taxa are the same as in the H-genome clade from Figure 3, but the analysis differs in that introns were excluded. Numbers above branches show ML bootstrap support ≥50%. “NA” and “Eu” distinguish North American and Eurasian Elymus species, respectively. Numbers following taxon names distinguish individuals within species where applicable, and are consistent among Figures 1–3, S1. Letters following these
improved the manuscript. Academic Editor Simon Joly provided an additional review, and we especially acknowledge his input regarding the complicated relationship between Hordum and Elymus.

Author Contributions
Conceived and designed the experiments: RMG. Performed the experiments: RMG MMB MN. Analyzed the data: RMG MMB MN. Contributed reagents/materials/analysis tools: RMG. Wrote the paper: RMG.

References
1. Guggisberg A, Bretagnolle F, Mansion G (2006) Allopolyploid origin of the Mediterranean endemic, Centaurium bionum (Gentianaceae), inferred by molecular markers. Systematic Botany 31: 368–379.
2. Hovese CE, Balsevich A (2005) Phylogenetic patterns and polyploid evolution within the Mediterranean centaurea Centaurea centaurium (Gentianaceae-Chiro- nieae). Taxon 54: 931–950.
3. Munnenhoff K, Franze A, Koch MA (1997) Molecular phylogenetics of Tilia s.l. (Tiliaceae) based on chloroplast DNA restriction site variation and sequences of the internal transcribed spacers of nuclear ribosomal DNA. Canadian Journal of Botany 75: 469–482.
4. Seeland T, Schnabel A, Wendel JF (1997) Congruence and consensus in the centaurea (Malvaceae). Systematic Botany 22: 259–260.
5. Widmer A, Balsbergber M (1999) Molecular evidence for allopolyploid speciation and a single origin of the narrow endemic Draba laudina (Brassicaceae). American Journal of Botany 86: 1292–1298.
6. Escobar García P, Schoenswetter P, Fuertes Aguilar J, Nieto Feliner G, 2002) Phylogenetic analysis of Eupersicaria indica. Molecular Phylogenetics and Evolution 22: 57–64.
7. Mason-Gamer RJ (2008) Allohexaploidy, introgression, and the complex relationship between perennial Triticeae (Poaceae). Journal of Systematics and Evolution 56: 467–478.
8. Arnheim N (1983) Concerted evolution of multigene families. In: Nei M, Yang W, eds. Molecular Evolution. New York: Oxford University Press. pp 315–364.
9. Hamby RK, Zimmer EA (1992) Ribosomal RNA as a phylogenetic tool in plant systematics. In: Soltis PS, Soltis DE, Doyle JJ, eds. Molecular Systematics of Plant Evolution. New York: Chapman and Hall. pp 267–280.
46. Lu B-R, von Bothmer R (1993) Mitotic analysis of Elymus canadensis, E. longisetus, and their interspecific hybrids with twenty-three Elymus species (Triticeae, Poaceae). Plant Systematics and Evolution 185: 35–53.
47. Salomon B (1993) Interspecific hybridizations in the Elymus semisteritus group (Poaceae). Genome 36: 809–805.
48. Salomon B, Lu B-R (1992) Genomic groups, morphology, and sectional delimitation in Eurasian Elymus (Poaceae, Triticeae). Plant Systematics and Evolution 180: 1–15.
49. Salomon B, Lu B-R (1994) Genomic relationships between species of the Elymus semisteritus group and Elymus semnus lato (Poaceae). Plant Systematics and Evolution 191: 199–211.
50. Deyweg DR (1973) The origin of Agropyron smithii. American Journal of Botany 62: 524–530.
51. Chen Q, Conner KL, Larche A, Thomas JB (1998) Genome analysis of Thionurus intermedium and Thionurus ponticum using genomic in situ hybridization. Genome 41: 580–586.
52. Liu Z-W, Wang RR-C (1993) Genome analysis of Elytrigia canescens, Lophopyrum nodosum, Psammoneuron geniculata sp. cynthis, and Thionurus intermedium (Triticeae, Gramineae). Genome 36: 102–111.
53. Zhang X-Y, Dong Y-S, Wang RR-C (1996) Characterization of genomes and chromosomes in partial amphiploids of the hybrid Tritium acaule x Thionurus ponticum by in situ hybridization, isozyme analysis, and RAPD. Genome 39: 1062–1071.
54. Deyweg DR (1982) Genomic and phylogenetic relationships among North American perennial Triticeae. In: Estes JR, ed. Grasses and grasslands. Norman: Oklahoma University Press. pp 51–81.
55. Mason-Gamer RJ (2007) Multiple homoplasious insertions and deletions of a GBSSI gene in Triticeae (Poaceae) allotetraploids: evidence from phosphoenolpyruvate carboxylase gene sequences. Systematic Botany 29: 850–861.
56. Davis JI, Echevarria C, Vidal J, Cejudo FJ (2002) Isolation and characterization of the wheat phosphoenolpyruvate carboxylase gene. Modelling of the encoded protein. Plant Science 162: 233–238.
57. Hellquist DM, Mason-Gamer RJ (2004) The evolution of North American Elymus (Triticeae, Poaceae) allotetraploids: evidence from phosphoenolpyruvate carboxylase gene sequences. Systematic Botany 29: 850–861.
58. Mason-Gamer RJ (2004) Reticulate evolution, introgression, and intertrigal inter gene capture in an allohexaploid grass. Systematic Biology 53: 25–37.
59. Mason-Gamer RJ (2005) The β-amylase genes of grasses and a phylogenetic analysis of the Triticeae (Poaceae). American Journal of Botany 92: 1045–1058.
60. Mason-Gamer RJ, Burns MM, Naum M (2010) Phylogenetic relationships and reticulation among Asian Elymus (Poaceae) allotetraploids: analysis of three nuclear genes. Molecular Phylogenetics and Evolution 54: 10–22.
61. Mason-Gamer RJ, Weil CF, Kellogg EA (1996) Granule-bound starch synthase: structure, function, and phylogenetic utility. Molecular Biology and Evolution 13: 1658–1673.
62. Rohde W, Becker D, Salamin F (1989) Structural analysis of the wax locus from Hordeum vulgare. Nuclear Acids Research 16: 7105–7106.
63. Wagler A, Blackstone N, Cartwright P, Dick M, Mose B, et al. (1994) Surveys of gene families using polymerase chain reaction: PCR selection and PCR drift. Systematic Biology 43: 250–261.
64. Bradley KD, Hillis DM (1993) Reticulatum DNA sequences generated by PCR amplification. Molecular Biology and Evolution 14: 592–593.
65. Cronn R, Cedroni M, Haselkorn T, Grover C, Wendel JF (2002) PCR-mediated recombination in amplification products derived from polyploid cotton. Theoretical and Applied Genetics 104: 492-499.
66. Judo M, Shelef AB, Wilson C (1990) Stimulation and suppression of PCR-mediated recombination. Nucleic Acids Research 26: 1819–1825.
67. Lepinec L, Kerey E, Philippe H, Gadal P, Cérin C (1993) Soghum phosphoenolpyruvate carboxylase gene family: structure, function, and molecular evolution. Plant Molecular Biology 21: 407–502.
68. Devos K, Bale MD (1997) Comparative genetics in the grasses. Plant Molecular Biology 35: 3–15.
69. Ziegler P (1999) Cereal beta-amylases. Journal of Cereal Science 29: 195–204.
70. Sharp P, Kreis M, Shevry P, Gale MD (1988) Location of β-amylase sequences in wheat and their relatives. Theoretical and Applied Genetics 75: 286–290.
71. Mason-Gamer RJ (2007) Multiple homoplasious insertions and deletions of a Triticeae (Poaceae) DNA transposon: a phylogenetic perspective. BMC Evolutionary Biology 7: 92.
72. Kleinholz A (1997) Integrating barley RFLP and classical marker maps. Barley Genetics Newsletter 27: 105–112.
73. Koruz V, Malychev S, Vovysenkov A, Boerner A (1997) RFLP-mapping of three mutant loci in rye (Secale cereale L) and their relation to homologous loci within the Gramineae. Theoretical and Applied Genetics 95: 468–473.
74. Frit F, Simon C, Sullivan J, Swofford DL (1997) Evolution of the mitochondrial cytochrome oxidase II gene in Collombola. Journal of Molecular Evolution 44: 145–158.
75. Sullivan J, Markert JA, Kilpatrick CW (1997) Phylogeny and molecular systematics of the Paspalum azetus species group (Redentia: Mucedae) inferred using parsimony and likelihood. Systematic Biology 46: 426–440.
76. Swofford DL, Olsen GJ, Waddell PJ, Hillis DM (1996) Phylogenetic inference. In: Hillis DM, Moritz C, Mable BK, eds. Molecular systematics, 2nd edition. Sunderland: Sinauer. pp 407–514.
77. Feulner J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. Journal of Molecular Evolution 17: 368–376.
78. Helsenbeck JP, Gratuil AA (1997) Phylogenetic estimation and hypothesis testing using maximum likelihood. Annual Review of Ecology and Systematics 28: 437–466.
79. Huelsenbeck JP, Rannala B (1997) Phylogenetic methods come of age: testing hypotheses in an evolutionary context. Science 276: 227–232.
80. Rodriguez F, Oliver JL, Marín A, Medrano JR (1990) The general stochastic model of nucleotide substitution. Journal of Theoretical Biology 142: 483–495.
81. Tavare S (1986) Some probabilistic and statistical problems in the analysis of DNA sequences. Lectures in Mathematics in the Life Sciences 17: 57–86.
82. Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro NH, ed. Mammalian protein metabolism. New York: Academic Press. pp 21–132.
83. Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16: 111–120.
84. Hasegawa M, Kishino H, Yano T (1983) Dating the human-ape split by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution 20: 502–510.
85. Tavare S (1986) Some probabilistic and statistical problems in the analysis of DNA sequences. Lectures in Mathematics in the Life Sciences 17: 57–86.
86. Hasegawa M, Kishino H, Yano T (1983) Dating the human-ape split by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution 20: 502–510.
87. Yang Z (1993) Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. Molecular Biology and Evolution 10: 1396–1401.
88. Zweidt 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.
89. Blattner FR (2004) Phylogenetic analysis of Hordeum (Poaceae) as inferred by nuclear rDNA ITS sequences. Molecular Phylogenetics and Evolution 33: 280–290.
90. Sun G-L, Ni Y, Daley T (2008) Molecular phylogeny of RPE2 gene reveals multiple origin, geographic distribution of H genome, and the relationship of the Y genome to other genomes of Elymus species. Molecular Phylogenetics and Evolution 46: 897–907.
91. Linde-Laursen I, Seberg O, Salomon B (1994) Comparison of the Giemsa C- and N-banded karyotypes of two Elymus species, E. dactylinus and E. glaucocen (Poaceae: Triticeae). Plant Systematics and Evolution 192: 163–176.
92. Jaaska V (1998) Isoenzyme data on the origin of North American allopolyploid Elymus species. In: Jaradat AA, ed. Triticeae III. Enfield, New Hampshire: Science Publishers, Inc. pp 209–216.
93. Jaaska V (1992) Isoenzyme variation in the grass genus Elymus (Poaceae). Hereditas 117: 11–22.
94. Tsujiimoto H, Gill BS (1991) Repetitive DNA sequences from polyploid Elymus trachycaulus and the diploid progenitor species: detection and genomic affinity of Elymus chromatin added to wheat. Genome 34: 782–789.