Supporting Information

Chester et al. 10.1073/pnas.1112041109

SI Materials and Methods

Seed Collections. Seeds were collected from natural populations, with one capitulum collected from each of 10 individuals. Between 10 and 15 seeds per individual were germinated and grown in a temperature-controlled greenhouse at the University of Florida. After 3 mo, one progeny per maternal plant from the field (a total of 60) and 10 sibling progeny (sharing at least the maternal parent and most likely both parents in this highly selfing species; ref. 1) from one individual from Spokane-1 (2729) were selected at random for analysis.

Collections were made over 3 y as follows: in 2009 for collections 2729, 2730, 2731, 2736, 2738; in 2010 for collections 2872 and 2875; and in 2011 for collection 2875-B. The collections were made in the following locations: 2729, West Sprague 3200 and South Ray Street, Spokane, WA; 2730, Appleway Boulevard and University Road, Spokane, WA; 2731, Sprague Avenue and Flora Road, Veradale, WA; 2736, Shetland Street near Mullan Avenue, Post Falls, ID; 2738, Sherman Avenue at freeway entrance to I-90, Coeur d'Alene, ID; 2872, Second Street and Pearl Street, Oakesdale, WA; 2875/2875-B, post office, South Grand Drive, Pullman, WA.

Progenitor DNA Repeat Identification/Isolation. Repetitive sequences were identified from genomic 454 sequences using an approach similar to that described previously (2). Genomic DNA (gDNA) was extracted from three *T. dubius* individuals (unique ID nos. 3937, 3927, and 3938) and two *T. pratensis* individuals (unique ID nos. 3920 and 3942) using a cetyltrimethylammonium bromide method (3). After RNA digestion with 20 μg of RNaseA (Fermentas) for 15 min at 37 °C, DNA was purified using the Qiagen DNeasy extraction kit. Approximately 5 μg of gDNA from each plant was sent to the Interdisciplinary Center for Biotechnology Research at the University of Florida for Genome Sequencer (GS) library construction and 454 sequencing (454 Life Sciences). Each sample was sequenced on one-eighth of a plate using a GS FLX instrument with GS FLX Titanium Series reagents (454 Life Sciences). From the resulting standard flowgram (.sff) files, a FASTA file of trimmed reads was generated for each sample using the sffinfo program (454 Life Sciences). Reads from the two *T. pratensis* samples were placed into one composite FASTA file, and all reads from *T. dubius* samples were placed in another FASTA file.

Each composite FASTA file was processed using the following procedure. First, sequences similar (E value $<1 \times 10^{-15}$) to four available Asteraceae plastid genome sequences (GenBank accession nos. NC_013553.1, NC_010601.1, NC_007977.1, and NC_007578.1) were identified using BLASTN (4). A new FASTA file was generated excluding reads matching the plastid sequences using a custom perl script that included bioperl modules (5). The resulting sequences yielded 158,792 (60,978,057 bp) and 233,541 (83,185,861 bp) plastid-filtered reads for *T. pratensis* and *T. dubius*, respectively, corresponding to estimated nuclear genome coverages of 1.9% for *T. pratensis* and 2.9% for *T. dubius*, using previously published genome sizes (6). For identifying tandemly repetitive sequences within each species, all reads per species were concatenated into one contiguous sequence, separated by Ns and processed using Tandem Repeats Finder (7). The output was organized using TRAP (8), and consensus repeat sequences were extracted into a new FASTA file. Genome abundance was estimated by conducting a BLASTN search for each monomer against the plastid-filtered gDNA 454 reads of both *Tragopogon* species. The resulting number of matches per repeat was extracted from the BLAST report and tabulated in Excel (Microsoft). Sequences that were highly represented in only *T. pratensis* or *T. dubius* were screened with Primer3 (9) in Geneious (Biomatters) for PCR primer design. The primers were used to amplify the repeat from genomic DNA of the *Tragopogon* species and cloned using standard methods. Two repeats, pra001 and dub005, were selected. Using either BLASTx or tBLASTx, neither sequence was found to have significant similarity (E value $<1 \times 10^{-15}$) to any sequence on GenBank. A single clone was checked by Sanger sequencing and used as a template for PCR amplification, nick translation labeling, and FISH, as described below.

Chromosome Preparation. The final 2 cm of growing roots were harvested and pretreated in an aqueous solution of 2 mM 8-hydroxyquinoline (Sigma-Aldrich) for 4.5 h at 4 °C. Pretreated roots were fixed in 3:1 ethanol-glacial acetic acid for 2 wk at −20 °C. Fixed roots were stored in 70% ethanol at −20 °C. Metaphase chromosome spreads were prepared as described previously (10).

FISH. All probes were labeled by nick translation in 5-μg batches following Birchler et al. (10). After labeling, all probes were purified using a QIAquick Nucleotide Removal Kit (Qiagen). FISH was performed using the protocol outlined by Birchler et al. (10) with a few minor modifications. For GISH, the hybridization mixture contained ∼450 ng of parental genomic DNA (gDNA) probe, 540 ng of autoclaved salmon sperm DNA in a final concentration of 0.7× sodium citrate (SSC) buffer (0.1 M NaCl and 0.01 M sodium citrate; pH 7.0). Probes were denatured on the slide in a covered tray, lined with moistened tissue, at 82–83 °C for 2.5 min. Slides were then transferred to a sealed humid chamber at 55 °C. Initially, slides were hybridized for 24 h; however, most plants were hybridized for 36–48 h, which led to an increased probe signal intensity. After hybridization, slides were washed briefly in 2× SSC at 55 °C and drained. A drop of Vectashield containing DAPI (Vector Laboratories) was added before mounting a glass coverslip (Corning).

Slides were viewed with a Zeiss Axio Imager.M2 fluorescence microscope, with fluorescence illumination provided by an X-Cite Series 120 Q Lamp (EXFO Life Sciences). Images were captured with a 100× objective lens and a microscope-mounted AxioCam MRm digital camera (Zeiss) in conjunction with Axiovision version 4.8 software (Zeiss) on a PC. Brightness and contrast for a captured image were adjusted in Axiovision by moving the upper and lower cutoffs in the histogram of signal intensity against the number of pixels. The Axiovision software was used to apply a color to the acquired gray-scale image as follows: DAPI was colored blue or gray, the *T. dubius* gDNA probe was colored green, and the *T. pratensis* gDNA probe was colored red. Axiovision was also used to merge different channels. All images were exported at 300 pixels per inch in .tif format into Adobe Photoshop CS3 version 10.0. In some cases, entire images were adjusted for brightness under the "levels" menu.

For FISH with repetitive DNA probes, clones were isolated from PCR products using a gDNA template of a *Tragopogon* species as indicated below. The subtelomeric and centromeric repeats correspond to TGP7 (Cy5 label, yellow) and TPRMBO (Cy3 label, red), respectively. These repeats, first described and characterized by Pires et al. (6), were used to design consensus primers for amplification as follows: TPRMBO_For, 5′-CACACCCTTGTGTGAAAAGGT-3′; TPRMBO_Rev, 5′-TTTCACGAAACTTCTTCAGTTAGC-3′; TGP7_For, 5′-TTGGCCC-

GTTTAAACTTCTG-3′; TGP7_Rev, 5′-CTTTACCACGTTG-TGCTCCA-3′. For the 35S rDNA probe, primers were designed using an existing 18S rDNA sequence for *T. dubius* (GenBank U42502.1): dub_18S_For, 5′-TGTGCCGGCGACGCATCATT-3′; dub_18S_Rev, 5′-GCGAGCTGATGACTCGCGCT-3′. For the 5S rDNA probe, PCR primers PI and PII were used, as described by Cox et al. (11). The dispersed genomic repeats pra001 and dub005 were amplified using the following consensus primers: pra_001_F, 5′-TGCGCTCCACAAACTCTTTCGAT-3′; pra_001_R, 5′-ATGGAGCGTAAACTTTTCGAACATGC-3′; dub_005_F, 5′-ACCCGAATTTCGAAACAGAACACCA-3′; and dub_005_R, 5′-TGGAAGTTCGGGGTCAAAAATGTAAGT-3′.

To generate FISH probes, a single representative plasmid clone, checked by Sanger sequencing, was used as a PCR template to generate DNA for nick translation labeling. The TGP7 clone comprised a 408-bp sequence amplified from *T. pratensis* (Cy5 label, yellow). The TPRMBO probe comprised a 470-bp sequence amplified from *T. porrifolius* (Cy3 label, red). The 35S rDNA clone comprised 1.3 kbp of the 18S rRNA-coding region amplified from *T. dubius* (fluorescein label, green). The 5S rDNA clone comprised a 266-bp sequence including the entire 5S rDNA intergenic spacer region amplified from *T. pratensis* (Cy3 label, red). The pra001 clone comprised a 174-bp sequence amplified from *T. pratensis* (Cy3 label, red). The dub005 clone comprised a 153-bp sequence amplified from *T. dubius* (fluorescein label, green).

The FISH experiments were carried out as described for GISH, but with 200–400 ng of each probe and a hybridization time of 16 h. After hybridization with FISH probes, chromosome preparations were GISH-reprobed in the following manner. The coverslip and mounting medium from the FISH procedure were removed by briefly washing the slide in 2× SSC at 55 °C. The slides were then air-dried at room temperature, UV cross-linked, and immediately hybridized with genomic DNA probes as described above.

Nuclear Microsatellites. Total genomic DNAs were used as template to amplify 12 progenitor-specific microsatellite loci from 65 *T. miscellus* individuals. These individuals were the same as those karyotyped with the exception of those from population 2738, where sibling progeny were used. Three plants, 2729–1, 2730–9, and 2738–1, were excluded because of missing data. Loci (1025, 1048, 1054, 1055, 1056, 1072, 1109, 1111d, 1111p, 1112d, 1112p, 1119) were amplified and sized following the methods of Symonds et al. (12). From these data, a pairwise genetic distance matrix was generated based on the proportion of shared alleles using Microsatellite Analyser (13). The distance matrix was used as input to produce a neighbor-net using SplitsTree (14).

1. Cook LM, Soltis PS (1999) Mating systems of diploid and allotetraploid populations of Tragopogon (Asteraceae), I: Natural populations. *Heredity* 82:237−244.
2. Macas J, Neumann P, Navrátilová A (2007) Repetitive DNA in the pea (*Pisum sativum* L.) genome: Comprehensive characterization using 454 sequencing and comparison to soybean and *Medicago truncatula*. *BMC Genomics* 8:427.
3. Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 19:11−15.
4. Altschul SF, et al. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25:3389−3402.
5. Stajich JE (2007) An introduction to BioPerl. *Methods in Molecular Biology*, ed Edwards D (Springer, Totowa, NJ), Vol 406, pp 535−548.
6. Pires JC, et al. (2004) Molecular cytogenetic analysis of recently evolved *Tragopogon* (Asteraceae) allopolyploids reveal a karyotype that is additive of the diploid progenitors. *Am J Bot* 91:1022−1035.
7. Benson G (1999) Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res* 27:573−580.
8. Sobreira TJP, Durham AM, Gruber A (2006) TRAP: Automated classification, quantification and annotation of tandemly repeated sequences. *Bioinformatics* 22:361−362.
9. Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 132:365−386.
10. Birchler JA, Albert SA, Gao Z (2008) Stability of repeated sequence clusters in hybrids of maize as revealed by FISH. *Trop Plant Biol* 1:34−39.
11. Cox AV, Bennett MD, Dyer TA (1992) Use of the polymerase chain reaction to detect spacer size heterogeneity in plant 5S ribosomal RNA gene clusters and to locate such clusters in wheat (*Triticum aestivum* L.). *Theor Appl Genet* 83:684−690.
12. Symonds VV, Soltis PS, Soltis DE (2010) Dynamics of polyploid formation in Tragopogon (Asteraceae): Recurrent formation, gene flow, and population structure. *Evolution* 64:1984−2003.
13. Dieringer D, Schlotterer C (2003) Microsatellite Analyser (MSA): A platform independent analysis tool for large microsatellite data sets. *Mol Ecol Notes* 3:167−169.
14. Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* 23:254−267.

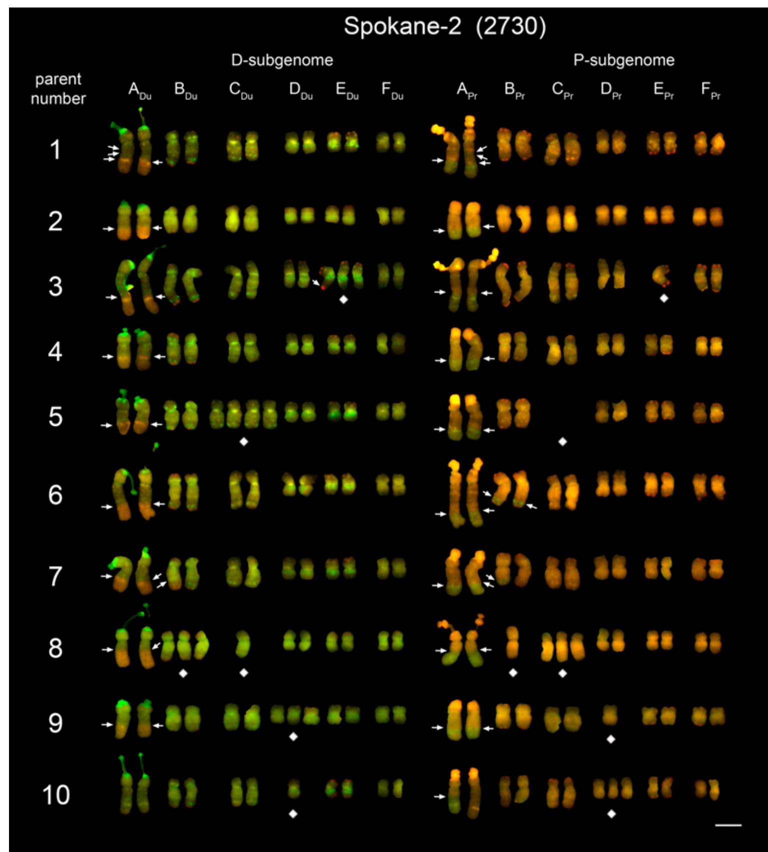

Fig. S1. GISH-probed mitotic karyotypes of *T. miscellus* progeny from Spokane-2, WA (2730). GISH was carried out with genomic DNA probes of *T. dubius* (green) and *T. pratensis* (red). Arrows indicate translocation breakpoints. Diamonds indicate aneuploid chromosomes. Plants 1–7 and 9 share a homozygous reciprocal translocation on A_{Du}/A_{Pr}. The translocation breakpoints for plant 8 (on A_{Du}/A_{Pr}) and plant 10 (on A_{Pr}) are positioned closer to the centromere. (Scale bar: 5 μm.)

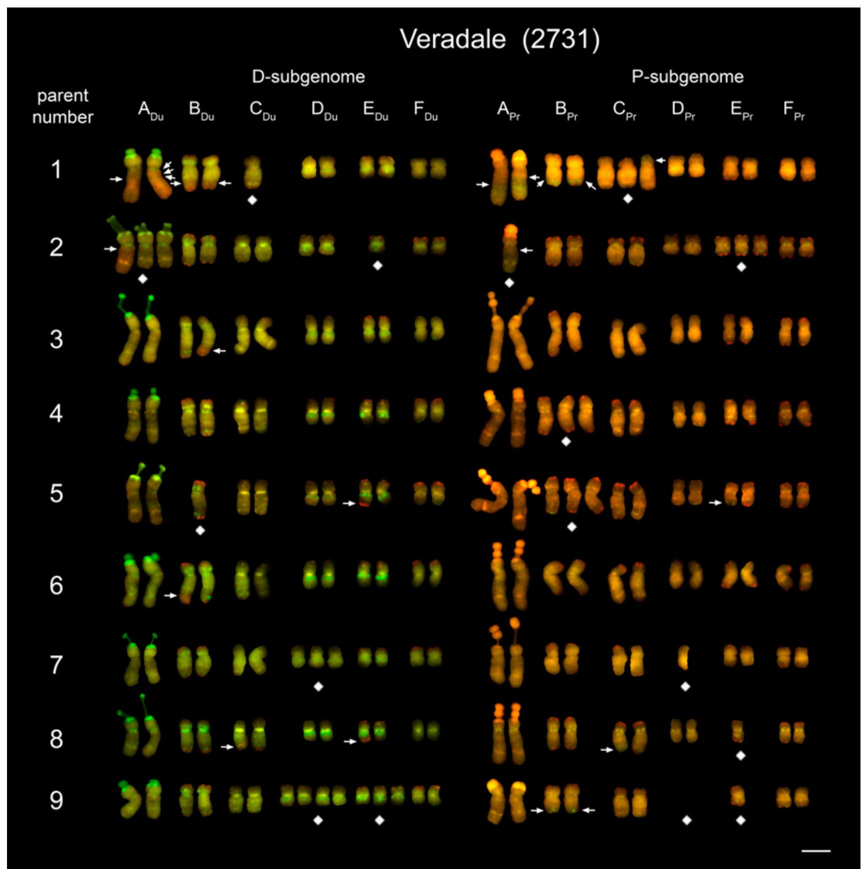

Fig. S2. GISH-probed mitotic karyotypes of *T. miscellus* progeny from Veradale, WA (2731). GISH was carried out with genomic DNA probes of *T. dubius* (green) and *T. pratensis* (red). Arrows indicate translocation breakpoints. Diamonds indicate aneuploid chromosomes. (Scale bar: 5 μm.)

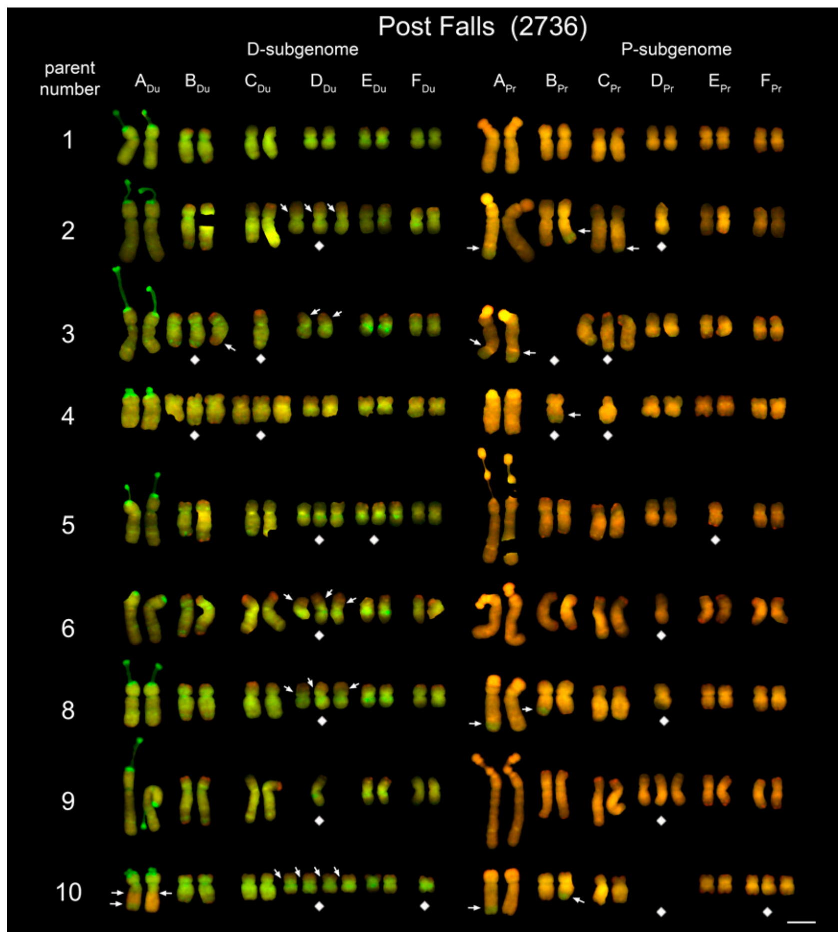

Fig. S3. GISH-probed mitotic karyotypes of *T. miscellus* progeny from Post Falls, ID (2736). GISH was carried out with genomic DNA probes of *T. dubius* (green) and *T. pratensis* (red). Arrows indicate translocation breakpoints. Diamonds indicate aneuploid chromosomes. (Scale bar: 5 μm.)

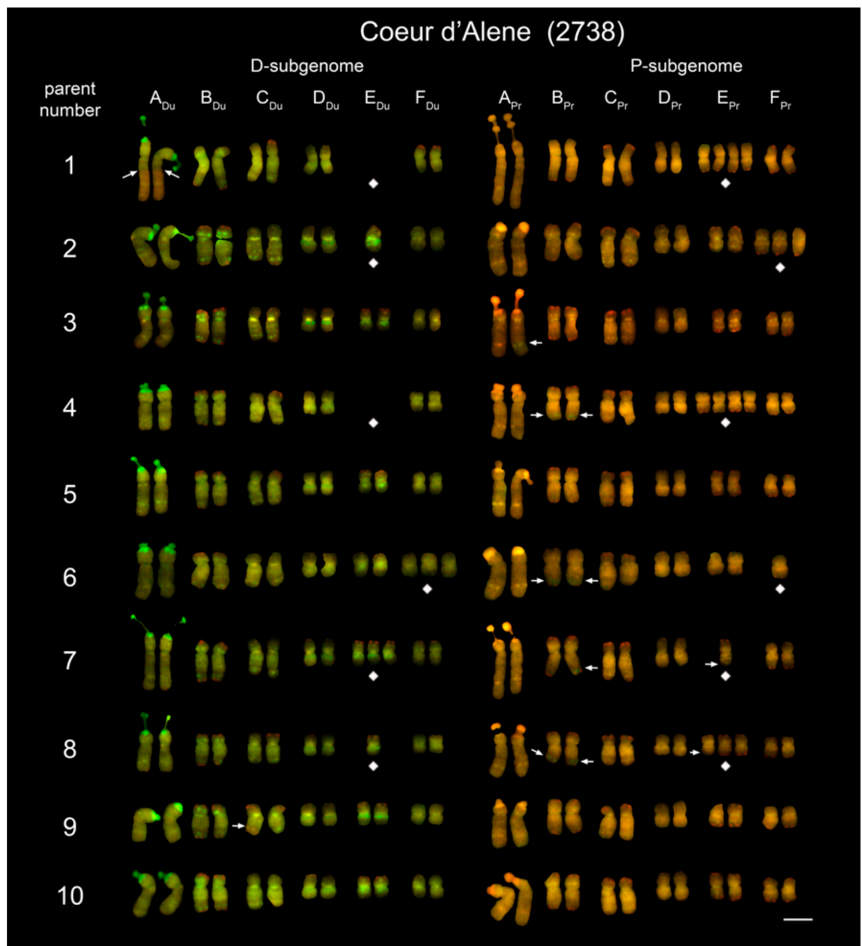

Fig. S4. GISH-probed mitotic karyotypes of *T. miscellus* progeny from Coeur d'Alene, ID (2738). GISH was carried out with genomic DNA probes of *T. dubius* (green) and *T. pratensis* (red). Arrows indicate translocation breakpoints. Diamonds indicate aneuploid chromosomes. (Scale bar: 5 μm.)

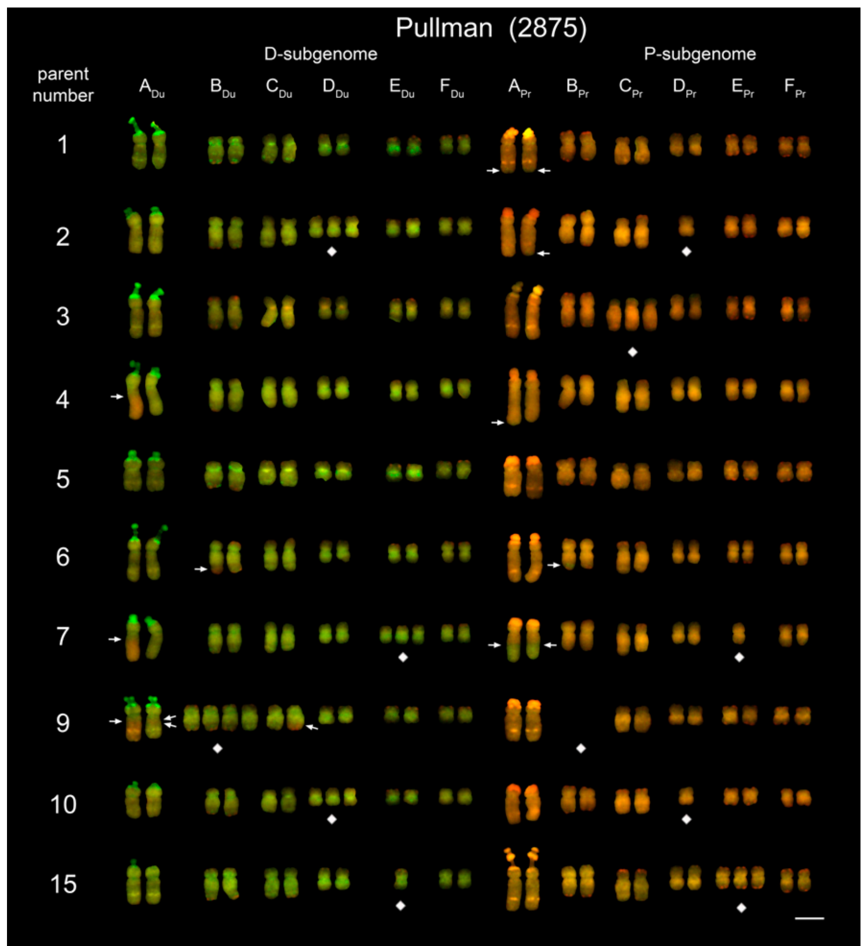

Fig. S5. GISH-probed mitotic karyotypes of *T. miscellus* progeny from Pullman, WA (2875). GISH was carried out with genomic DNA probes of *T. dubius* (green) and *T. pratensis* (red). Arrows indicate translocation breakpoints. Diamonds indicate aneuploid chromosomes. (Scale bar: 5 μm.)

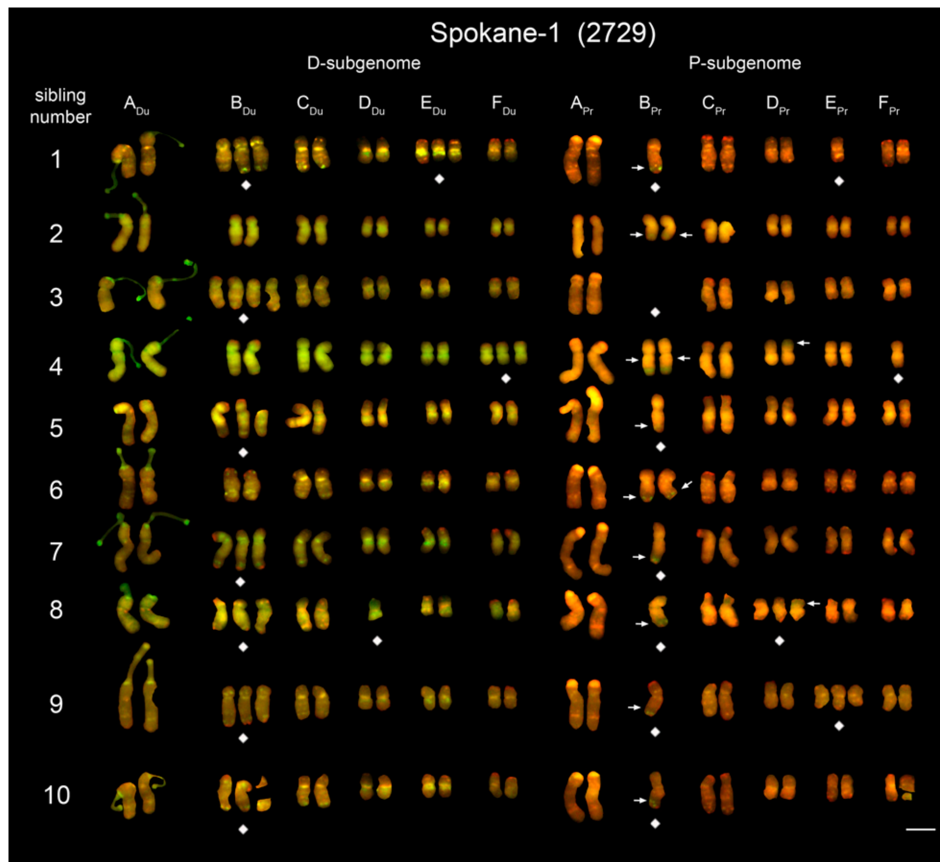

Fig. S6. GISH-probed mitotic karyotypes of *T. miscellus* sibling progeny from Spokane-1, WA (2729). GISH was carried out with genomic DNA probes of *T. dubius* (green) and *T. pratensis* (red). Arrows indicate translocation breakpoints. Diamonds indicate aneuploid chromosomes. (Scale bar: 5 μm.)

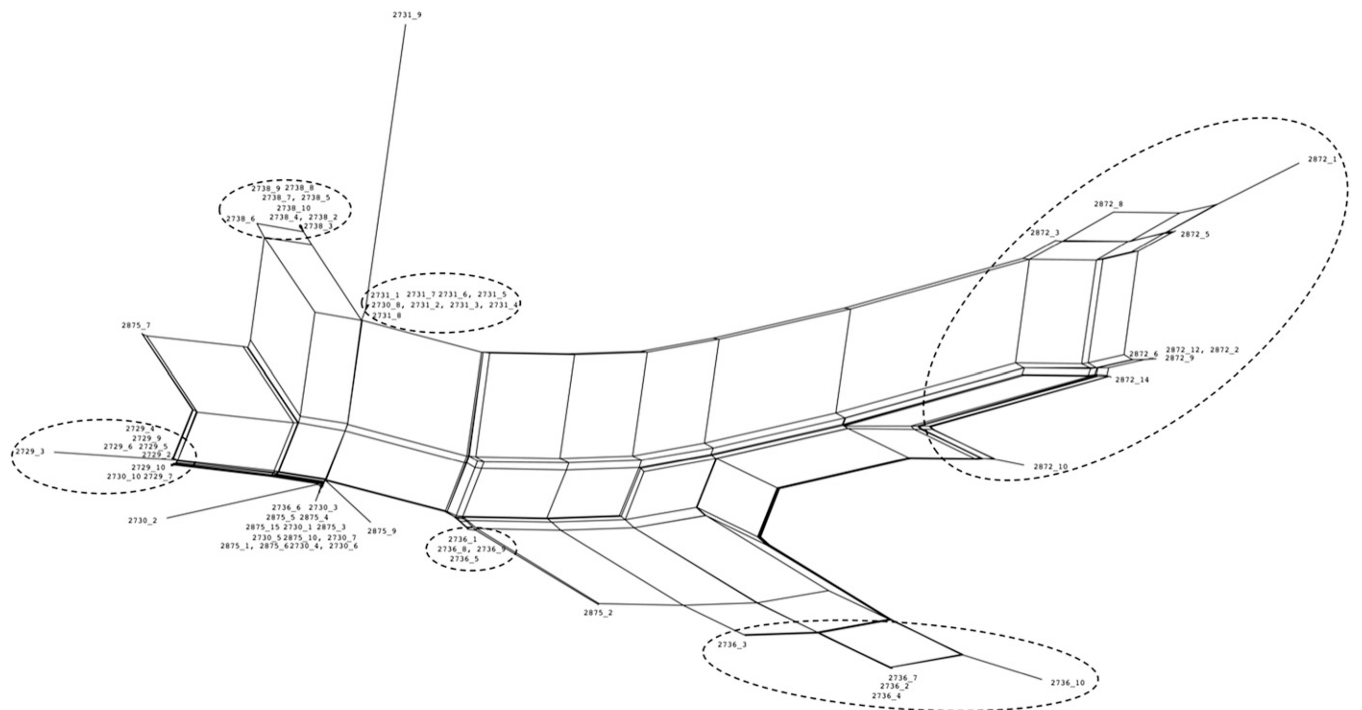

Fig. S7. Neighbor-net of *T. miscellus* individuals from seven populations. Dashed ovals enclose clusters of individuals from the same population; in some instances, an individual from another population may occur within a cluster. The results indicate that each sampled population represents an independent origin of *T. miscellus*; population 2736 includes individuals from two distinct origins. Although most individuals from populations 2730 and 2875 cluster together in the neighbor-net, other data show that the two populations formed independently (e.g., 2875 is the only *T. miscellus* population with *T. dubius* as maternal parent). Some individuals did not cluster with other members of their population (e.g., 2730–8-3 and 2730–10-1), possibly because of mutation and/or interpopulation crossing.

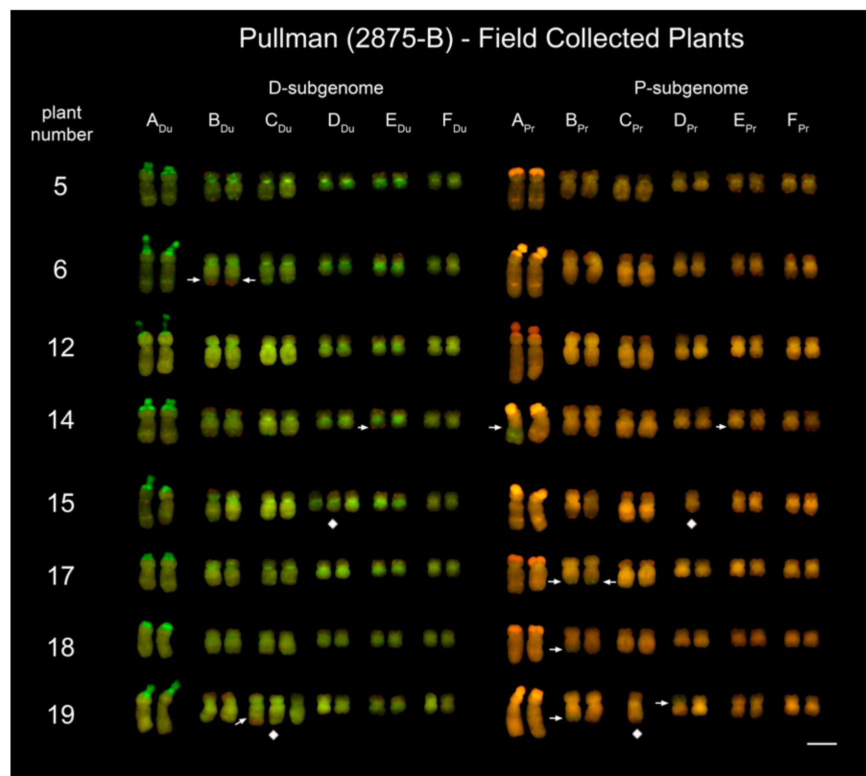

Fig. S8. GISH-probed mitotic karyotypes of *T. miscellus* adult plants growing in the town of Pullman, WA. GISH was carried out with genomic DNA probes of *T. dubius* (green) and *T. pratensis* (red). Arrows indicate translocation breakpoints. Diamonds indicate aneuploid chromosomes. (Scale bar: 5 μm.)

Table S1. Karyotype summary of all *T. miscellus* plants analyzed

Plant no.	Chromosome no.	Monosomy/<u>nullisomy</u>*	Trisomy/<u>tetrasomy</u>*	Reciprocal translocation(s)†	Nonreciprocal translocation(s)†
2730–1-1	24			A	
2730–2-1	24			**A**	
2730–3-1	24	E_{Pr}	E_{Du}	**A**	E_{Du}
2730–4-1	24			**A**	
2730–5-1	24	$\underline{C_{Pr}}$	$\underline{C_{Du}}$	**A**	
2730–6-1	24			**A**	$\mathbf{B_{Pr}}$
2730–7-1	24			**A** B	
2730–8-3	24	B_{Pr} C_{Du}	B_{Du} C_{Pr}	**A**	
2730–9-1	24	D_{Pr}	D_{Du}	**A**	
2730–10-1	24	D_{Du}	D_{Pr}		A_{Pr}
2731–1-1	24	C_{Du}	C_{Pr}	A **B**	C_{Pr}
2731–2-1	24	A_{Pr} E_{Du}	A_{Du} E_{Pr}	A	
2731–3-6	24				B_{Du}
2731–4-1	25		B_{Pr}		
2731–5-15	24	B_{Du}	B_{Pr}	E	
2731–6-1	24				B_{Du}
2731–7-1	24	D_{Pr}	D_{Du}		
2731–8-1	23	E_{Pr}		C	E_{Du}
2731–9-1	24	$\underline{D_{Pr}}$ E_{Pr}	$\underline{D_{Du}}$ E_{Du}		$\mathbf{B_{Pr}}$
2731–10-1	ns	—	—	—	—
2736–1-1	24				
2736–2-1	24	D_{Pr}	D_{Du}		A_{Pr} B_{Pr} $\mathbf{D_{Du}}$
2736–3-1	23	$\underline{B_{Pr}}$ C_{Du}	B_{Du} C_{Pr}		$\mathbf{A_{Pr}}$ B_{Du} $\mathbf{D_{Du}}$
2736–4-1	24	B_{Pr} C_{Pr}	B_{Du} C_{Du}		$\mathbf{B_{Pr}}$
2736–5-1	25	E_{Pr}	D_{Du} E_{Du}		
2736–6-1	24	D_{Pr}	D_{Du}		$\mathbf{D_{Du}}$
2736–7-1	ns	—	—	—	—
2736–8-1	24	D_{Pr}	D_{Du}		A_{Pr} B_{Pr} $\mathbf{D_{Du}}$
2736–9-2	24	D_{Du}	D_{Pr}		
2736–10-1	24	$\underline{D_{Pr}}$ F_{Du}	$\underline{D_{Du}}$ F_{Pr}		A_{Pr} A_{Du} B_{Pr} $\mathbf{D_{Du}}$
2738–1-2	24	$\underline{E_{Du}}$	$\underline{E_{Pr}}$		$\mathbf{A_{Du}}$
2738–2-1	24	E_{Du}	F_{Pr}		
2738–3-1	24				A_{Pr}
2738–4-2	24	$\underline{E_{Du}}$	$\underline{E_{Pr}}$		$\mathbf{B_{Pr}}$
2738–5-1	24				
2738–6-1	24	F_{Pr}	F_{Du}		$\mathbf{B_{Pr}}$
2738–7-1	24	E_{Pr}	E_{Du}		B_{Pr} $\mathbf{E_{Pr}}$
2738–8-1	24	E_{Du}	E_{Pr}		$\mathbf{B_{Pr}}$ E_{Pr}
2738–9-1	24				C_{Du}
2738–10-1	24				
2872–1-1	24	A_{Pr} $\underline{D_{Pr}}$	A_{Du} $\underline{D_{Du}}$		A_{Du}
2872–2-1	24	$\underline{E_{Pr}}$	$\underline{E_{Du}}$		A_{Du} A_{Pr}
2872–3-1	24	E_{Pr} D_{Du}	D_{Pr} E_{Du}		B_{Du}
2872–5-1	24	B_{Du} C_{Du} E_{Du} D_{Pr}	B_{Pr} C_{Pr} E_{Pr} D_{Du}		B_{Pr} E_{Pr}
2872–6-1	24				$\mathbf{E_{Du}}$
2872–8-1	24	E_{Pr}	E_{Du}		$\mathbf{D_{Du}}$ C_{Pr}
2872–9-1	24	A_{Pr} $\underline{B_{Du}}$ D_{Du}	A_{Du} $\underline{B_{Pr}}$ D_{Pr}	E	A_{Du}
2872–10-1	24				$\mathbf{A_{Du}}$ A_{Pr}
2872–12-1	24	B_{Du}	B_{Pr}		
2872–14-1	25		C_{Pr}		C_{Pr}
2875–1-1	24				$\mathbf{A_{Pr}}$
2875–2-1	24	D_{Pr}	D_{Du}		A_{Pr}
2875–3-1	25		C_{Pr}		
2875–4-1	24				A_{Du} A_{Pr}
2875–5-1	24				
2875–6-1	24			B	
2875–7-1	24	E_{Pr}	E_{Du}		A_{Du} $\mathbf{A_{Pr}}$
2875–9-1	24	$\underline{B_{Pr}}$	$\underline{B_{Du}}$		A_{Du} C_{Du}
2875–10-1	24	D_{Pr}	D_{Du}		
2875–15-1	24	E_{Du}	E_{Pr}		
2729–2-1	24	B_{Pr} E_{Pr}	B_{Du} E_{Du}		$\mathbf{B_{Pr}}$
2729–2-2	24				$\mathbf{B_{Pr}}$
2729–2-3	24	$\underline{B_{Pr}}$	$\underline{B_{Du}}$		

Table S1. Cont.

Plant no.	Chromosome no.	Monosomy/nullisomy*	Trisomy/tetrasomy*	Reciprocal translocation(s)[†]	Nonreciprocal translocation(s)[†]
2729–2-4	24	F_{Pr}	F_{Du}		$\mathbf{B}_{\mathbf{Pr}}$ D_{Pr}
2729–2-5	24	B_{Pr}	B_{Du}		$\mathbf{B}_{\mathbf{Pr}}$
2729–2-6	24				$\mathbf{B}_{\mathbf{Pr}}$
2729–2-7	24	B_{Pr}	B_{Du}		$\mathbf{B}_{\mathbf{Pr}}$
2729–2-8	24	B_{Pr} D_{Du}	B_{Du} D_{Pr}		$\mathbf{B}_{\mathbf{Pr}}$ D_{Pr}
2729–2-9	25	B_{Pr}	B_{Du} E_{Pr}		$\mathbf{B}_{\mathbf{Pr}}$
2729–2-10	24	B_{Pr}	B_{Du}		$\mathbf{B}_{\mathbf{Pr}}$
2875-B-5	24				
2875-B-6	24				$\mathbf{B}_{\mathbf{Du}}$
2875-B-12	24				
2875-B-14	24			E	A_{Pr}
2875-B-15	24	D_{Pr}	D_{Du}		
2875-B-17	24				$\mathbf{B}_{\mathbf{Pr}}$
2875-B-18	24				B_{Pr}
2875-B-19	24	C_{Pr}	C_{Du}		B_{Pr} C_{Du} D_{Pr}

*An upper-case letter designates the chromosome, and subscript letters indicate parental origin based on the GISH signal at and around the centromere.
[†]Homologs that were fully homozygous in terms of translocation(s) are shown in bold type.