Transposable Element Proliferation and Genome Expansion Are Rare in Contemporary Sunflower Hybrid Populations Despite Widespread Transcriptional Activity of LTR Retrotransposons

Takeshi Kawakami[1], Preeti Dhakal[1], Angela N. Katterhenry[1], Chelsea A. Heatherington[2], and Mark C. Ungerer*[,1]

[1]Division of Biology, Kansas State University

[2]Department of Forest Resources and Conservation, University of Florida

*Corresponding author: E-mail: mcungere@ksu.edu.

We sequenced a small number of RT-PCR bands (coding regions of LTR retrotransposons) for confirmation purposes. These sequences have been deposited in GenBank under accession numbers HQ665469–HQ665504.

Accepted: 24 January 2011

Abstract

Hybridization is a natural phenomenon that has been linked in several organismal groups to transposable element derepression and copy number amplification. A noteworthy example involves three diploid annual sunflower species from North America that have arisen via ancient hybridization between the same two parental taxa, *Helianthus annuus* and *H. petiolaris*. The genomes of the hybrid species have undergone large-scale increases in genome size attributable to long terminal repeat (LTR) retrotransposon proliferation. The parental species that gave rise to the hybrid taxa are widely distributed, often sympatric, and contemporary hybridization between them is common. Natural *H. annuus* × *H. petiolaris* hybrid populations likely served as source populations from which the hybrid species arose and, as such, represent excellent natural experiments for examining the potential role of hybridization in transposable element derepression and proliferation in this group. In the current report, we examine multiple *H. annuus* × *H. petiolaris* hybrid populations for evidence of genome expansion, LTR retrotransposon copy number increases, and LTR retrotransposon transcriptional activity. We demonstrate that genome expansion and LTR retrotransposon proliferation are rare in contemporary hybrid populations, despite independent proliferation events that took place in the genomes of the ancient hybrid species. Interestingly, LTR retrotransposon lineages that proliferated in the hybrid species genomes remain transcriptionally active in hybrid and nonhybrid genotypes across the entire sampling area. The finding of transcriptional activity but not copy number increases in hybrid genotypes suggests that proliferation and genome expansion in contemporary hybrid populations may be mitigated by posttranscriptional mechanisms of repression.

Key words: hybridization, transposable elements, derepression, genome evolution, repetitive DNA.

Introduction

Transposable elements are DNA sequences that can mobilize and replicate within the genome of their host. Two major classes of elements are recognized based on the nature of the transposition intermediate and mechanism of mobilization (Wicker et al. 2007). Class I elements transpose via an RNA intermediate transcribed from an existing element; intermediates are reverse transcribed prior to insertion at a new location. The class II elements transpose via a DNA intermediate that is either excised from its existing location via double-stranded cleavage followed by subsequent insertion elsewhere or by a phenomenon involving transposition in which only a single strand is cleaved (e.g., rolling circle replication, [Kapitonov and Jurka 2001]). The vast amounts of genome sequence data now available for a large and growing number of organisms coupled with rapid advances in analytical approaches for detecting repetitive elements

have greatly facilitated studies of transposable element biology (Saha et al. 2008; Feschotte et al. 2009; Pritham 2009; Lerat 2010). Consequently, our understanding of these sequences and their potential roles in shaping genome evolution has grown significantly in recent years (Pritham 2009; Tenaillon et al. 2010).

Class I elements are particularly relevant in evolutionary analyses of genome size and structure because of their proliferative mode of replication. Autonomous long terminal repeat (LTR) retrotransposons are the most abundant Class I elements in plants and are subdivided into two superfamilies known as *Ty3/Gypsy*-like and *Ty1/Copia*-like (Kumar and Bennetzen 1999). These superfamilies display similar structural features (e.g., flanking LTRs in direct orientation, presence of *GAG* and *POL* genes) but represent separate ancient lineages diverged at the sequence level and that differ with regard to domain order within *POL*. Despite their proliferative capacity, transcriptional and/or transpositional activity of most Class I elements is repressed by the host genome through a combination of epigenetic mechanisms involving both transcriptional and posttranscriptional controls (Zilberman and Henikoff 2004; Slotkin and Martienssen 2007; Lisch 2009). Because of the highly mutagenic consequences of unchecked transposition, historical selection likely has been strong for the maintenance of these host repression mechanisms. Under certain genomic and environmental conditions, however, host repression may fail, resulting in large-scale and episodic activation and proliferation (Wessler 1996; Grandbastien 1998; Lisch 2009).

Hybridization is a naturally occurring phenomenon shown in some species to be associated with derepression of LTR retrotransposons (Waugh O'Neill et al. 1998; Labrador et al. 1999; Metcalfe et al. 2007; Lisch 2009; Michalak 2009). The molecular basis of hybridization-induced element derepression is not fully understood but likely involves disruption of one or more host repression systems as a consequence of merging differentiated genomes. A potential example of such derepression involves three annual sunflower species from North America. These species (*Helianthus anomalus*, *H. deserticola*, *and H. paradoxus*) arose via ancient hybridization events between the same two extant sunflower parental species, *H. annuus* and *H. petiolaris*. The hybrid species independently have undergone massive proliferation events of *Ty3/gypsy*-like elements and, to a lesser extent, *Ty1/copia*-like elements (Ungerer et al. 2006, 2009; Staton et al. 2009; Kawakami et al. 2010). These proliferation events occurred in recent evolutionary time, as evidenced both by the young ages of the hybrid species (Schwarzbach and Rieseberg 2002; Welch and Rieseberg 2002; Gross et al. 2003) and by estimates of proliferation events themselves (Ungerer et al. 2009).

Because of their broad geographic and often sympatric distributions, natural hybridization between *H. annuus* and *H. petiolaris* remains an ongoing and common phenomenon in wild populations. Contemporary *H. annuus* × *H. petiolaris* hybrid populations typically consist of a mixture of genotypes ranging from early generation hybrid to advanced backcross individuals and display a broad array of hybrid index scores (Rieseberg et al. 1998, 1999). Natural zones of hybridization likely served as sources for the establishment and evolution of the hybrid species and, as such, represent excellent natural laboratories to examine genome dynamics of hybridization, including the potential role of hybridization in transposable element derepression and proliferation.

In the current report, we examine whether natural hybridization between *H. annuus* × *H. petiolaris* may be a catalyst for transposable element derepression and genome expansion. We demonstrate that, despite massive proliferation events in the genomes of all three sunflower hybrid species, natural hybridization between the two parental species is not, in-and-of-itself, a major trigger of transposable element proliferation, and that genome size increases in hybrid genotypes are rare. Interestingly, however, we detect widespread transcriptional activity of *Ty3/gypsy* and *Ty1/copia* sequences in both hybrid and nonhybrid genotypes. These results suggest posttranscriptional mechanisms of repression may be preventing further copy number amplification of these sequences in wild sunflower populations.

Materials and Methods

Plant Materials

During the summer of 2008, seeds were collected from naturally occurring *H. annuus* × *H. petiolaris* hybrid populations and nearby (within ~1 km) stands of "pure" *H. annuus* and *H. petiolaris* from several locations in the central, western, and southwestern United States (table 1). At two locations (Keith Co., Nebraska and Quay Co., New Mexico), two or three nearby but distinct hybrid populations were identified and sampled. The two hybrid populations from Keith Co. were the focus of previous studies by Rieseberg et al. (1998, 1999). The remaining hybrid populations were newly identified for the current study.

At each hybrid and parental population, seeds were collected from mature capitula on 20~30 plants in the field. Seeds were transported to Kansas State University and germinated in the dark on moist filter paper in Petri dishes. Seedlings were transferred to 4-inch pots and allowed to grow in the Kansas State University greenhouses until suitable size for the harvesting and processing of leaf tissue for 1) genome size determination via flow cytometry and 2) DNA extraction for estimation of transposable element numerical abundance via quantitative polymerase chain reaction (PCR) and microsatellite analyses of hybridity. DNA was extracted using a DNeasy Plant Mini Kit (Qiagen) following the manufacturer's instructions. Sample sizes for the different experimental assays of hybrid and nonhybrid genotypes are indicated in table 1.

Table 1
Location of Populations Used in This Experiment and the Number of Individuals Sampled in Various Assays

County, State	Population		N_{FC}	$N_{Q\text{-}PCR}$	$N_{RT\text{-}PCR}$	Latitude (N), Longitude (W)
El Paso Co., Colorado	ElP-A	*Helianthus annuus*	30	8	2	38.8679, 104.6832
	ElP-HP1	Hybrid population	27	10	4	38.9116, 104.6276
	ElP-P	*H. petiolaris*	24	8	2	38.9115, 104.6031
Keith Co., Nebraska	Kei-A	*H. annuus*	20	9	2	41.3268, 101.7258
	Kei-HP1	Hybrid population	30	8	4	41.2507, 101.6828
	Kei-HP2	Hybrid population	29	6	—	41.1638, 101.6890
	Kei-P	*H. petiolaris*	24	10	2	41.2578, 101.6942
Sandoval Co., New Mexico	San-A	*H. annuus*	29	10	2	36.1474, 107.2641
	San-HP1	Hybrid population	26	9	4	36.1098, 107.1767
	San-P	*H. petiolaris*	25	8	2	36.1861, 107.3717
Quay Co., New Mexico	Qu-A	*H. annuus*	25	7	2	35.1104, 103.2942
	Qu-HP1	Hybrid population	27	8	4	35.1266, 103.2440
	Qu-HP2	Hybrid population	27	9	—	35.1140, 103.2611
	Qu-HP3	Hybrid population	26	9	—	35.1140, 103.2479
	Qu-P	*H. petiolaris*	30	8	2	35.1551, 103.2434
Barton Co., Kansas	Bar-A	*H. annuus*	25	11	2	38.4155, 98.6568
	Bar-HP1	Hybrid population	30	20	3	38.3552, 98.6636
	Bar-P	*H. petiolaris*	27	11	2	38.3414, 98.6641

NOTE.—N_{FC}, Number of individuals for which estimates of genome size were determined by flow cytometry; $N_{Q\text{-}PCR}$, number of individuals for which estimates of *Ty3/gypsy* and *Ty1/copia* abundance were determined by Q-PCR. These individuals also were genotyped at 22 SSR loci to determine hybrid index scores; $N_{RT\text{-}PCR}$, number of individuals sampled in transcriptional assays of *Ty3/gypsy* and *Ty1/copia* sequences.

Microsatellite Assay

Natural *H. annuus* × *H. petiolaris* hybrid populations can be identified phenotypically as they consist of individuals displaying a range of phenotypes spanning the variability across the two species, with many individuals showing intermediate characteristics. However, to 1) confirm these hybrid populations as such and 2) examine aspect of genome size and LTR retrotransposon copy number variation in relation to hybrid index scores, we genotyped subsets of individuals from hybrid populations and nearby stands of parental species at 22 previously developed microsatellite loci (supplementary table S1, Supplementary Material online). For the microsatellite assays, we utilized an M13-tailed primer protocol (Schuelke 2000) with four possible fluorophores (NED, FAM, VIC, and PET). PCR reactions were performed in 15 μl volumes consisting of 1× PCR buffer (Promega), 2.5 mM $MgCl_2$, 200 μM of each dNTP, 50 nM of a forward M13-tailed primer, 50 nM of a M13 dye-labeled primer (5′-ACGACGTTGTAAAACGAC), 100 nM reverse primer, 1 unit of Taq polymerase, and 2–12 ng of template DNA.

PCR was conducted on an MJ Research PTC-100 thermal cycler using the following touchdown procedure: initial denaturing at 95 °C for 5 min; followed by 5 cycles of 95 °C for 45 s, 68 °C for 5 min (−2 °C/cycle), 72 °C for 1 min; followed by 5 cycles of 95 °C for 45 s, 58 °C for 2 min (−2 °C/cycle), 72 °C for 1 min; followed by 25 cycles of 95 °C for 45 s, 50 °C for 2 min, 72 °C for 1 min; and a final extension of 72 °C for 5 min. PCR was conducted separately for each locus but with amplification products of 3–4 loci (each utilizing a different dye) subsequently pooled for co-loading on an ABI 3730 DNA Analyzer. Raw data were scored using GeneMarker software Version 1.85 (SoftGenetics). Hybrid index scores were determined for each hybrid population using the program HINDEX V 1.42 (Buerkle 2005) with parental allele frequencies estimated from nearby stands of *H. annuus* and *H. petiolaris* populations.

Genome Size Determination

Nuclear DNA content (2C genome size) was estimated using a FACSCalibur flow cytometer (Becton Dickinson). A minimum of 10,000 nuclei was analyzed per sample. To prepare samples for assay, fresh leaves (50 mg) were chopped with razor blades in a buffer containing 15 mM N-2-hydroxyethylpiperazine-N′-2-ethanesulfonic acid, 1 mM ethylenediaminetetraacetic acid, 80 mM KCl, 20 mM NaCl, 300 mM sucrose, 0.2% Triton-X, 0.5 mM spermine, 0.1% b-mercaptoethanol (modified after Bino et al. 1993), filtered through 30-mm nylon mesh, and then centrifuged to collect nuclei. The samples were stained for 2 h with 700 ml of intercalating propidium iodide solution (BioSure) and 2 ml of an internal standard solution containing chicken erythrocyte nuclei (CEN; BioSure). Duplicate analyses were conducted for a subset of individuals from hybrid populations (5∼6 individuals per population) to determine the repeatability and precision of the estimates.

Because considerable variation can exist in reported 2C values of CEN among studies (reviewed in Bennett and Leitch 1995 and references therein), we estimated 2C value of CEN by using *H. deserticola* (one of the sunflower hybrid species) as an external standard. Leaves from *H. deserticola*

(n = 4 plants, Ames 26094; United States Department of Agriculture [USDA], National Plant Germplasm System [NPGS]) were analyzed with the method described above. The analysis was duplicated on two different days. Mean peak positions of CEN relative to *H. deserticola* nuclei (2C = 10.79 pg, [Baack et al. 2005]) gave an estimated 2C value of CEN = 2.689 pg (standard deviation [SD] = 0.152). To validate that this 2C value of CEN yields appropriate genome size estimates of samples, 2C genome size of *H. anomalus* (Ames 26095; USDA, NPGS) also was estimated. Our estimate of 2C genome size of *H. anomalus* (2C = 11.03∼11.83, n = 2) was consistent with a previous report for this species (2C = 11.46 pg, [Baack et al. 2005]). The appropriateness of the 2C value of CEN in our experiment was thus confirmed. The same CEN batch was utilized for genome size estimates of all samples reported in this study.

LTR Retrotransposon Copy Number Estimation

Copy number estimates of specific lineages of *Ty3/gypsy* and *Ty1/copia* LTR retrotransposons in the genomes of hybrid and pure parental individuals were estimated via quantitative PCR using a standard curve method described in Ungerer et al. (2006) and Kawakami et al. (2010). Briefly, a 168 bp *Ty3/gypsy* integrase (Int) fragment and a 230 bp *Ty1/copia* RNase H fragment were generated by standard PCR using genomic DNA of *H. annuus* with the following primers: forward, 5′-TTATGCATTGCTTGCCAAAG-3′, and reverse, 5′-TCGACTCACCAAGTCTGCAC-3′ for *Ty3/gypsy* Int domain, and forward, 5′-TCTCAGAACCTCGGCAATCT-3′, and reverse, 5′-GGCGAGCAAAAGAGAAAATG-3′, for *Ty1/copia* RNase H domain. These fragments were cloned using a StrataClone PCR Cloning Kit (Agilent Technologies). The concentration of isolated plasmids (and insert) was determined by a NanoDrop spectrophotometer (Thermo Scientific), and plasmid number per unit volume was estimated. A 5-fold dilution series of the cloned fragments was utilized (50, 10, 2, 0.4, and 0.08 pg/μl) to generate standard curves that were used to estimate absolute abundances of elements in our samples.

Quantitative PCR assays were performed using 1 μl of cloned DNA for standard curves or sample DNA (standardized to 80 pg/μl) with an iCycler iQ™ quantitative PCR system (Bio-Rad). Each reaction was performed in duplicate using the iQ™ SYBR Green Supermix kit (Bio-Rad) following the manufacturers protocols, with the exception that reactions were conducted in 25 μl volumes. PCR efficiency for the standard curve exceeded 96%. Copy number estimates were converted to copy number per genome using mean *C*-values for *H. annuus* and *H. petiolaris* estimated from several populations (Baack et al. 2005). The average *C*-value for *H. annuus* and *H. petiolaris* (6.96 pg) was used for hybrid individuals. Assays of LTR retrotransposon copy number were conducted on the same set of individuals that were genotyped with microsatellite markers, with the exception of individuals from the Barton Co. Kansas sampling location.

Statistical Analysis

Estimated nuclear DNA content and copy number estimates of *Ty3/gypsy*-like and *Ty1/copia*-like sequences were compared by analysis of variance (ANOVA). Particular hybrid individuals exhibiting elevated nuclear DNA content or high copy number estimates of *Ty3/gypsy* or *Ty1/copia* elements were subjected to the Dixon outlier detection test (Verma and Quiroz-Ruiz 2006). Correlation between nuclear DNA content and copy numbers of *Ty3/gypsy*-like or *Ty1/copia*-like sequences was examined by Pearson product-moment correlations. All statistical analyses were performed in the software JMP 7.0.1 (SAS institute).

Transcriptional Assays

Transcriptional assays of *Ty3/gypsy* and *Ty1/copia* elements were conducted using the AccessQuick reverse transcriptase (RT)-PCR System (Promega). The same primer pairs used to estimate genomic copy numbers of *Ty3/gypsy* and *Ty1/copia* elements (see above) were utilized in assays of transcriptional activity of these retrotransposons. Primers targeting the 18s rRNA subunit (forward, 5′-CTTCGGGATCGGAGTAATGA-3′, and reverse, 5′-CTAAGAAGCTGGCCATGGAG-3′) were used for positive control reactions. To test for genomic DNA contamination of RNA, negative control reactions were performed by withholding the AMV RT. RT-PCR amplification was carried out by incubating at 45 °C for 45 min for the reverse transcription step, followed by one cycle of 95 °C for 2 min, 22 cycles (94 °C for 30 s; 56 °C for 30 s; and 72 °C for 60 s), and a final incubation cycle at 72 °C for 5 min. The amplified products were size-separated via electrophoresis in 2% agarose gels and visualized with ethidium bromide. Total RNA was extracted using TRIzol (Invitrogen) and purified with a RNeasy Plant Mini Kit (Qiagen) following the manufacturer's instructions. RNA was treated with RNase-Free DNase (Qiagen) to eliminate genomic DNA contamination.

RT-PCR amplified cDNA fragments of *Ty3/gypsy* and *Ty1/copia* were cloned from one *H. annuus*, *H. petiolaris*, and hybrid individual from the Barton County, KS location using a StrataClone PCR Cloning Kit. From 7 to 10 clones, per element per individual were sequenced on an ABI 3730XL DNA Analyzer (Applied Biosystems). DNA polymorphisms and sequence diversity indices were estimated using DnaSP ver. 5 (Librado and Rozas 2009).

Results

Genome Size Estimates in Hybrid and Nonhybrid Genotypes

At each of the five regional sampling locations (table 1), individuals from *H. annuus* populations consistently exhibit

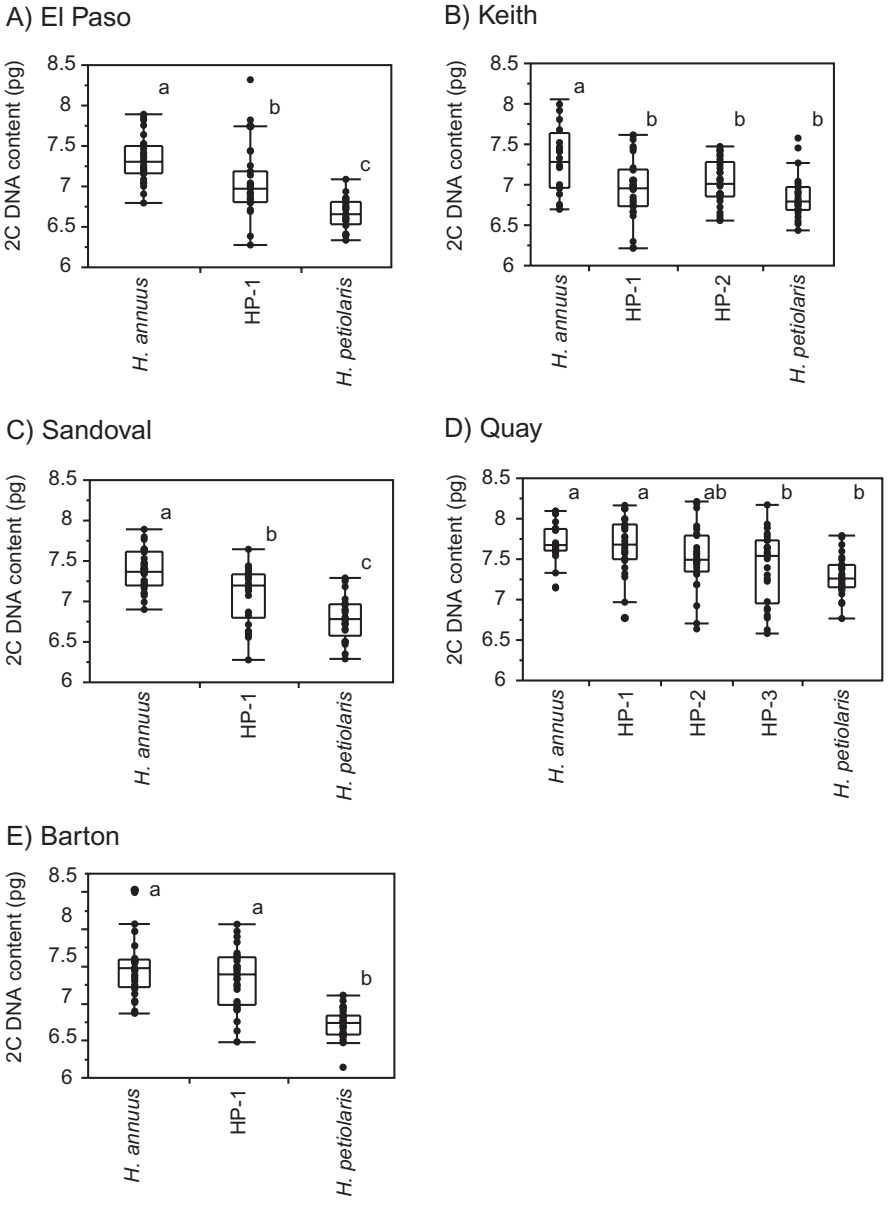

Fig. 1.—Estimates of nuclear DNA content (2C genome size) for individuals in natural *Helianthus annuus* × *H. petiolaris* hybrid populations (HP-x) and nearby stands of pure *H. annuus* and *H. petiolaris*. Data are shown for five different sampling locations (subpanels A–E). Lower case letters indicate results of Turkey–Kramer HSD post hoc tests. Number of individuals sampled: $n = 129$ for *H. annuus*, $n = 130$ for *H. petiolaris*, and $n = 222$ for individuals from hybrid populations.

higher mean 2C values than those from *H. petiolaris* populations (7.31 ± 0.424 – 7.69 ± 0.247 for *H. annuus* vs. 6.66 ± 0.192 – 7.29 ± 0.234 for *H. petiolaris*) (fig. 1; table 2). At those same locations, individuals from hybrid populations display 2C values largely intermediate between those observed for *H. annuus* and *H. petiolaris* (fig. 1, table 2), indicating that large-scale genome expansion in individuals from hybrid populations is rare. Significant differences in 2C values among *H. annuus*, *H. petiolaris*, and the hybrid populations were found at all sampling locations (table 2). Variance in 2C values typically was greater in hybrid populations than in populations of parental species, suggesting that a diversity of hybrid genotypes likely exists within the hybrid populations sampled. Several individuals from the Quay County, NM, hybrid population and a single individual from the El Paso County, CO, hybrid population exhibit 2C values that exceed values observed for either parental species. These values, however, were only slightly higher than the maximum values observed for *H. annuus* individuals. The Dixon outlier detection test (Verma and Quiroz-Ruiz 2006) failed to identify these values as significant outliers ($P > 0.05$).

Table 2
Estimates of Nuclear DNA Content and Copy Numbers of *Ty3/gypsy* and *Ty1/copia* Sequences in Individuals from *Helianthus annuus*, *H. petiolaris*, and Natural Hybrid Populations

			Nuclear DNA Content (pg/2C Genome)				*Ty3/gypsy* Copy Number				*Ty1/copia* Copy Number			
County	**Population**		**Mean**	**SD**	***F***	***P***	**Mean**	**SD**	***F***	***P***	**Mean**	**SD**	***F***	***P***
El Paso Co., Colorado	ElP-A	*H. annuus*	7.32	0.274	30.141	<0.0001	53,369	3,609	34.541	<0.0001	57,134	3,915	26.517	<0.0001
	ElP-HP1	Hybrids	7.04	0.426			26,019	3,228			30,292	3,502		
	ElP-P	*H. petioralis*	6.66	0.192			9,804	3,609			25,131	3,915		
Keith Co., Nebraska	Kei-A	*H. annuus*	7.31	0.424	7.326	0.0002	62,376	3,403	31.213	<0.0001	53,261	3,691	11.029	<0.0001
	Kei-HP1	Hybrids	6.95	0.376			13,457	3,609			23,957	3,915		
	Kei-HP2	Hybrids	7.03	0.275			35,495	4,168			23,422	4,521		
	Kei-P	*H. petioralis*	6.85	0.276			12,187	3,228			29,565	3,502		
Sandoval Co., New Mexico	San-A	*H. annuus*	7.38	0.253	27.182	<0.0001	37,664	3,228	27.163	<0.0001	21,590	3,691	22.646	<0.0001
	San-HP1	Hybrids	7.09	0.345			24,454	3,403			21,889	3,691		
	San-P	*H. petioralis*	6.79	0.277			12,249	3,609			14,626	3,915		
Quay Co., New Mexico	Qu-A	*H. annuus*	7.69	0.247	7.365	<0.0001	34,288	3,858	6.382	0.0005	36,446	3,502	3.354	0.0197
	Qu-HP1	Hybrids	7.67	0.345			22,658	3,609			23,914	3,691		
	Qu-HP2	Hybrids	7.50	0.387			24,177	3,403			14,135	3,915		
	Qu-HP3	Hybrids	7.39	0.435			16,784	3,403			35,302	4,186		
	Qu-P	*H. petioralis*	7.29	0.234			7,243	3,609			21,356	3,915		
Barton Co., Kansas	Bar-A	*H. annuus*	7.50	0.425	35.55	<0.0001	39,492	3,078	39.923	<0.0001	56,944	3,339	11.677	0.0001
	Bar-HP1	Hybrids	7.33	0.398			32,176	2,283			46,991	2,476		
	Bar-P	*H. petioralis*	6.72	0.202			8,259	3,078			30,929	3,339		

Within parental species, mean 2C values were significantly different among the five regional sampling locations (*H. annuus:* ANOVA, $F_{4,124} = 5.98$, $P = 0.0002$; *H. petiolaris:* ANOVA, $F_{4,125} = 31.39$, $P < 0.0001$). Mean 2C values also were significantly different among hybrid populations across the sampling area (ANOVA, $F_{7,214} = 13.44$, $P < 0.0001$). Populations of *H. annuus* and *H. petiolaris* from Quay County, NM, displayed the highest species-specific mean 2C values. Correspondingly, hybrid populations from this location also displayed higher 2C values than hybrid populations from other locations.

LTR Retrotransposon Copy Number Estimates in Hybrid and Nonhybrid Genotypes

Because flow cytometry may lack sensitivity to detect genome size changes associated with smaller-scale increases in LTR retrotransposon copy numbers, we performed quantitative PCR assays of *Ty3/gypsy* and *Ty1/copia* copy number abundance in a subset of hybrid individuals and nonhybrid individuals across the five regional collection locations. In previous reports (Ungerer et al. 2006; Kawakami et al. 2010), we identified specific sublineages of *Ty3/gypsy* and *Ty1/copia* that have undergone proliferation events in the hybrid sunflower species. Copy number estimates of these same lineages were determined for 6–20 individuals from each hybrid population and nearby stands of pure *H. annuus* and *H. petiolaris*. In accordance with observations of overall genome size (fig. 1), *H. annuus* individuals displayed a higher average abundance of *Ty3/gypsy* and *Ty1/copia* sequences compared with *H. petiolaris* (fig. 2, table 2), with estimates for hybrid populations again generally intermediate. A single hybrid individual from the Barton County, KS, population was identified with a copy number estimate of *Ty3/gypsy* that exceeded those observed for either parental species; however, this individual was not identified as an outlier (Dixon outlier test, $P > 0.05$). None of the hybrid individuals with slightly or moderately elevated 2C values (see fig. 1*A* and *D*) had correspondingly elevated copy number estimates of *Ty3/gypsy* and/or *Ty1/copia* sequences. Population-level differences in copy numbers of *Ty3/gypsy* and *Ty1/copia* sequences were found for *H. annuus* (*Ty3/gypsy*: ANOVA, $F_{4,40} = 8.02$, $P < 0.0001$; *Ty1/copia:* ANOVA, $F_{4,40} = 4.72$, $P = 0.0032$), *H. petiolaris* (*Ty3/gypsy:* ANOVA, $F_{4,40} = 6.29$, $P = 0.0005$; *Ty1/copia*: ANOVA, $F_{4,40} = 10.60$, $P < 0.0001$), and for the hybrid populations (*Ty3/gypsy:* ANOVA, $F_{7,71} = 3.74$, $P = 0.0017$; *Ty1/copia:* ANOVA, $F_{7,71} = 11.77$, $P < 0.0001$).

LTR Retrotransposon Copy Number Estimates and Hybrid Index Scores

Given differences in copy number abundance of retrotransposons between *H. annuus* and *H. petiolaris* (fig. 2) and given the lack of evidence of copy number amplification in individuals from hybrid populations (figs. 1 and 2), copy number estimates of *Ty3/gypsy* and *Ty1/copia* sequences in individuals from hybrid populations are expected to be correlated with hybrid index scores, which estimate the relative proportion of *H. petiolaris* versus *H. annuus* DNA in the genome. Figure 3 depicts the relationship between copy number abundance and hybrid index score for *Ty3/gypsy* (fig. 3*A*) and *Ty1/copia*

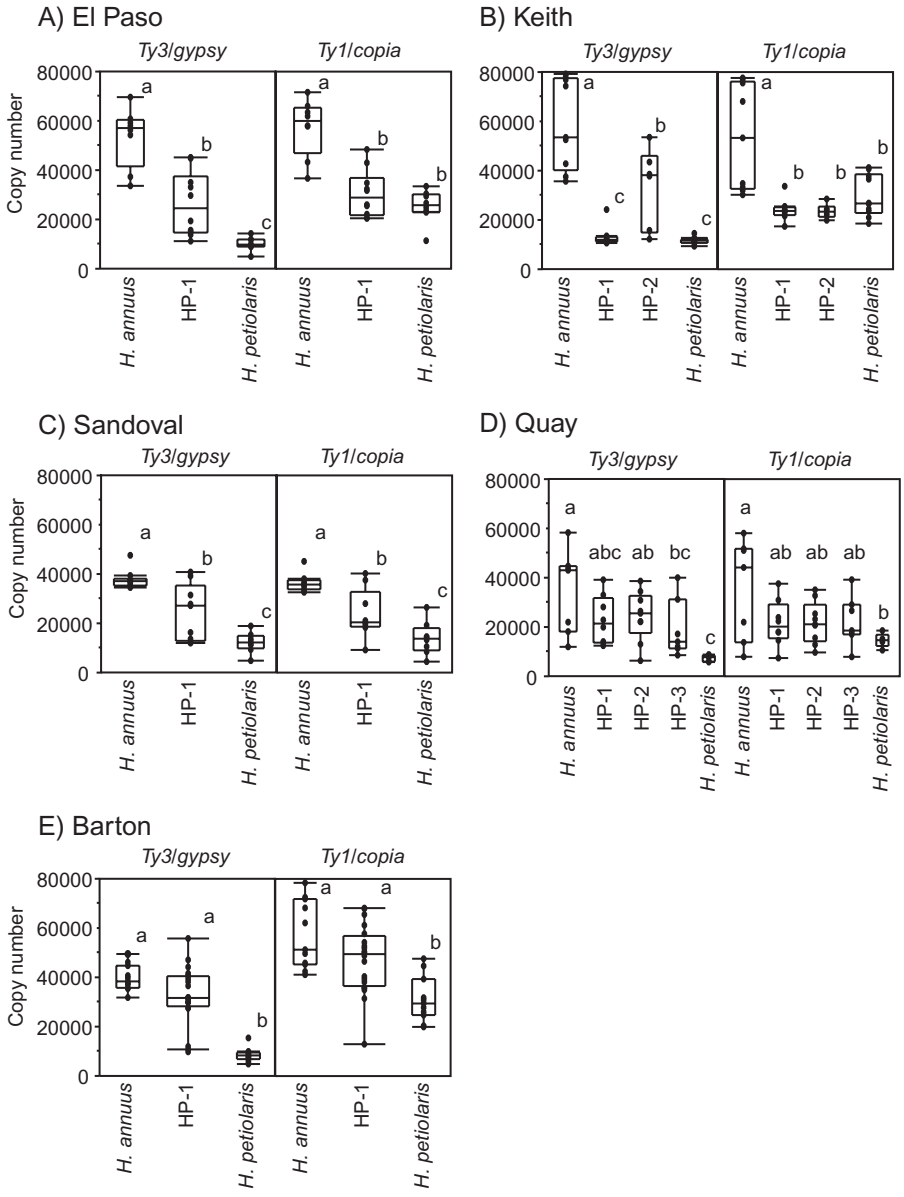

Fig. 2.—Estimates of genome copy number abundance of *Ty3/gypsy*-like and *Ty1/copia*-like sequences for individuals in natural *Helianthus annuus* × *H. petiolaris* hybrid populations (HP-x) and nearby stands of pure *H. annuus* and *H. petiolaris*. Data are shown for five different sampling locations (subpanels A–E). Lower case letters indicate results of Turkey–Kramer HSD post hoc tests. Number of individuals sampled: $n = 45$ for *H. annuus*, $n = 45$ for *H. petiolaris*, and $n = 79$ for individuals from hybrid populations.

(fig. 3*B*) elements where hybrid index values of 0 and 1 represent *H. petiolaris*-like and *H. annuus*-like genotypes, respectively. Consistent with expectations, higher copy number estimates are associated with hybrid index scores closer to 1 (more *H. annuus*-like) for both superfamilies, although this relationship is stronger for *Ty3/gypsy* (fig. 3).

LTR Retrotransposon Expression Assays

Transcriptional activity of *Ty3/gypsy* and *Ty1/copia* elements was assayed via RT-PCR for a total of 10 *H. annuus* individuals, 10 *H. petiolaris* individuals, and 19 individuals from hybrid populations across the five sampling locations. These assays were performed using the same primer pairs utilized to determine element copy number abundance via quantitative PCR (see Materials and Methods). Transcripts were detected for both types of element in all hybrid individuals as well as in all assayed individuals of *H. annuus* and *H. petiolaris* (fig. 4). DNA contamination was ruled out by negative control reactions in which the RT enzyme was withheld (fig. 4). Amplicons were confirmed as *Ty3/gypsy* and *Ty1/copia*

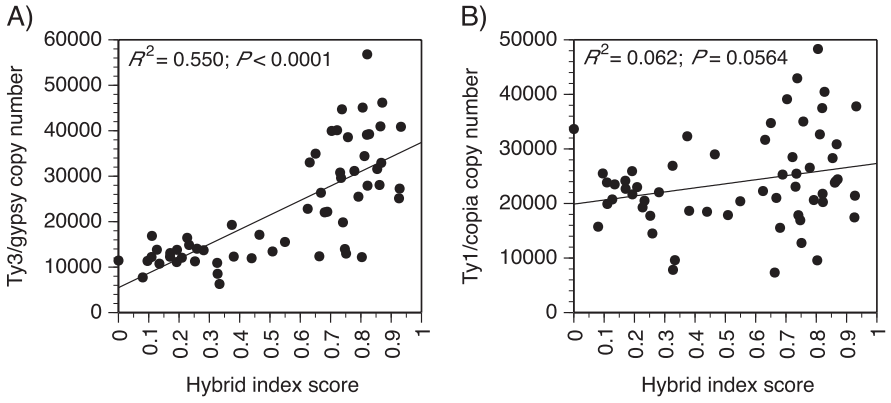

Fig. 3.—Genomic copy number estimates of *Ty3/gypsy* (*A*) and *Ty1/copia* (*B*) elements as a function of hybrid index score for 59 hybrid individuals from four different regional sampling locations. Hybrid index scores range from 0 (*H. petiolaris*-like) to 1 (*Helianthus annuus*-like). Individuals from the Barton Co. Kansas site were not included in this analysis (see Materials and Methods).

transcripts by cloning and sequencing PCR products for one *H. annuus*, one *H. petiolaris*, and one hybrid individual from the Barton Co, KS location. Between 7 and 10 clones were sequenced per amplicon per individual. Sequenced transcripts exhibited polymorphism within and between individuals (table 3), with higher levels of polymorphism observed for *Ty3/gypsy*. These patterns are consistent with a previous report documenting transcriptional activity of *Ty3/gypsy* and *Ty1/copia* elements in an inbred line of *H. annuus* (Vukich et al. 2009). All sequences have been deposited in GenBank under accession numbers HQ665469–HQ665504.

Discussion

Hybridization has been linked to activation of transposable elements in several organismal groups, with most examples documenting such activity in early generation hybrid genotypes (Waugh O'Neill et al. 1998; Labrador et al. 1999; Shan et al. 2005; Petit et al. 2010). In previous reports, we described the occurrence and dynamics of historical proliferation events of LTR retrotransposons in the genomes of three sunflower species that independently arose via hybridization between the same two parental taxa (Ungerer et al. 2006, 2009; Staton et al. 2009; Kawakami et al. 2010). Temporal estimates of the origins of the hybrid species are variable (Schwarzbach and Rieseberg 2002; Welch and Rieseberg 2002; Gross et al. 2003; Ungerer et al. 2009), but all estimates suggest their formation within the last 1 My. Although bouts of transposable element proliferation in these species genomes are clearly associated with hybridization followed by genome reorganization (i.e., hybrid speciation), the extent to which natural hybridization in-and-of-itself was, and may continue to be, a catalyst for large-scale changes in genome size is unclear. The parental species, *H. annuus* and *H. petiolaris*, which gave rise to the hybrid taxa, have broad and overlapping geographic distributions, and interspecific hybridization between them is common (Rieseberg et al. 1999). Natural *H. annuus* × *H. petiolaris* hybrid populations likely served as original source populations from which the three hybrid species lineages became established and, as such, represent excellent natural laboratories for investigating the direct role of hybridization on transposable element dynamics in this group.

In the current report, genome size was estimated for a total of 222 individuals from hybrid populations and 259 individuals from nearby populations of pure parental species (n = 129 for *H. annuus*; n = 130 for *H. petiolaris*) over multiple locations where these species occur in sympatry. Consistent with previous reports (Baack et al. 2005), genome size estimates for *H. annuus* were consistently higher than those for *H. petiolaris*. These interspecific differences in genome size were mirrored by differences in copy number abundance of both *Ty3/gypsy* and *Ty1/copia* sequences, suggesting that differences in genome size between these species are attributable, at least in part, to differential abundance of LTR retrotransposons. Genome size estimates of individuals from hybrid populations were generally intermediate between those observed for *H. annuus* and *H. petiolaris*, suggesting that hybridization is not a major trigger of transposable element derepression in natural populations.

A small number of individuals from hybrid populations displayed genome size values that exceeded those measured for either parental species (fig. 1). These increases were small, however, especially compared with those that took place historically in the sunflower hybrid species, *H. anomalus*, *H. deserticola*, and *H. paradoxus*, where 2C values are at least 50% greater than those estimated for either parental species (Baack et al. 2005). Hybrid individuals exhibiting slightly elevated genome size values typically displayed hybrid index scores closer to 1 (more *annuus*-like) and did not demonstrate elevated copy number estimates

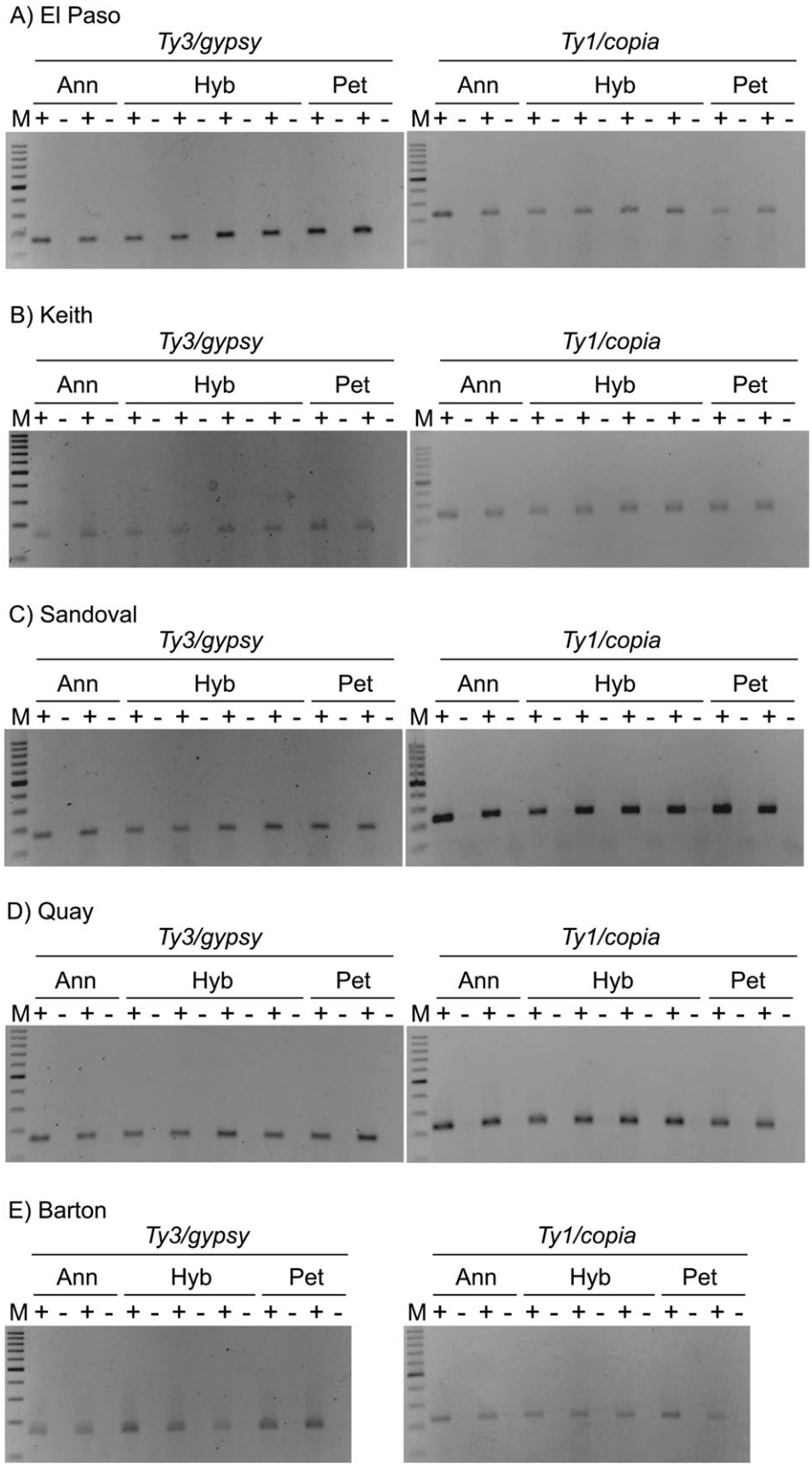

Fig. 4.—Transcriptional assays of *Ty3/gypsy* and *Ty1/copia* elements in *Helianthus annuus*, *H. petiolaris*, and hybrid individuals collected from five different regional sampling locations (subpanels A–E). Plus (+) and minus (–) signs indicate whether RT enzyme was provided or withheld from the RT-PCR reaction, respectively. Ann, *H. annuus*; Pet, *H. petiolaris*; Hyb, hybrid individual

Table 3
Nucleotide Variation of *Ty3/gypsy* and *Ty1/copia* Transcript Sequences Isolated from a Single Individual of *Helianthus annuus*, *H. petiolaris*, and a Natural Hybrid Genotype from the Barton County, KS Sampling Location

	Ty3/gypsy				*Ty1/copia*			
	N^a	S^b	h^c	π^d	N^a	S^b	h^c	π^d
H. annuus	7	29	4	0.115	8	19	8	0.029
H. petiolaris	8	41	8	0.120	10	15	7	0.027
Hybrids	7	30	7	0.114	8	2	4	0.005
All individuals	22	53	19	0.136	26	29	17	0.024

NOTE.—Sequences of *Ty3/gypsy* Integrase and *Ty1/copia* RNase H fragments have been deposited in GenBank (accession numbers HQ665469–HQ665504).

[a] The number of sequences obtained.

[b] The number of segregating sites.

[c] The number of unique sequences.

[d] Sequence diversity.

of LTR retrotransposons in direct assays of numerical abundance via quantitative PCR. Hybrid index score thus proved the best predictor of LTR retrotransposon abundance (especially for *Ty3/gypsy* elements) and genome size in hybrid genotypes. Taken together, these observations suggest that proliferation of transposable element sequences in natural sunflower hybrid populations is indeed rare. These findings are consistent with several recent studies demonstrating an absence of transposable element mobilization in both laboratory generated (Beaulieu et al. 2009) and naturally occurring (Kentner et al. 2003; Hazzouri et al. 2008; Parisod et al. 2009) interspecific plant hybrids.

The lack of evidence for genome size expansion and LTR retrotransposon copy number increases across these five regional locations of *H. annuus* × *H. petiolaris* hybridization is notable and contrasts sharply with the scale and repeatability of proliferation events that took place in the genomes of the ancient hybrid species. It should be noted, however, that assays of LTR retrotransposon abundance via quantitative PCR, whereas appropriate for detecting large-scale proliferation, are likely to miss smaller-scale mobilization events that could be documented by other methods, such as transposon display (Waugh et al. 1997; Van den Broeck et al. 1998; Melayah et al. 2001). The quantitative PCR assay provides reasonable resolution, however, as determined by variation between the two technical replicates performed for each individual in the current study. For example, across the 169 samples assayed by this method (table 1), pairs of technical replicates differed on average by 8.4% (SD = 7.1%) and 10.3% (SD = 8.4%) for assays of *Ty3/gypsy* and *Ty1/copia* elements, respectively.

Although the goal of the current study was to evaluate natural hybridization as a possible catalyst for transposable element proliferation, several lines of evidence suggest that biotic and abiotic stressors also may facilitate activation of transposable element sequences (Wessler 1996; Grandbastien 1998; Lisch 2009). Interestingly, the three hybrid species whose genomes underwent proliferation are locally adapted to relatively harsh environmental conditions. *Helianthus anomalus* and *H. deserticola* occupy desert regions, whereas *H. paradoxus* inhabits saline marshes (Rieseberg et al. 2003; Gross et al. 2004; Donovan et al. 2010). It thus seems feasible that some combination of hybridization and abiotic stress may have been required for the proliferation events that took place in the hybrid taxa. Experiments testing these ideas have been initiated in the greenhouse where environmental stress can be more precisely controlled.

Although proliferation of LTR retrotransposons leading to genome size increases was not detected in this study, transcriptional assays indicate that lineages of *Ty3/gypsy* and *Ty1/copia* elements that proliferated in the hybrid species remain active transcriptionally in hybrid populations as well as in natural populations of *H. annuus* and *H. petiolaris* across the entire sampling area. Sequencing of a small number of transcripts from *H. annuus*, *H. petiolaris*, and hybrid individuals from one of the sampling locations indicates that multiple *Ty3/gypsy* and *Ty1/copia* variants are transcribed. Transcriptional activity of *Ty3/gypsy* and *Ty1/copia* lineages was recently documented in an inbred line of *H. annuus* (Vukich et al. 2009) as well as in greenhouse-generated *H. annuus* × *H. petiolaris* hybrids and the parental lines used to construct them (B. Hartman-Bakken, Kawakami, Ungerer, unpublished data). The *Ty3/gypsy* and *Ty1/copia* transcripts isolated from *H. annuus*, *H. petiolaris*, and hybrid individuals bear strong resemblance to elements previously characterized in sunflower via sequence surveys (Ungerer et al. 2009; Kawakami et al. 2010) or shown previously to be transcriptionally active in *H. annuus* (Vukich et al. 2009). In addition, Blast searches against a *Helianthus* expressed sequence tag database with these transcripts as queries identified highly similar sequences isolated from multiple *Helianthus* species, suggesting widespread transcriptional activity of these sublineages in sunflowers.

If, as our data indicate, transcriptional activity of these elements is commonplace and widespread, the lack of insertional activity and genome expansion in contemporary hybrids is intriguing and suggests mechanisms of transposable element repression acting posttranscriptionally, such as RNAi (Slotkin and Martienssen 2007; Lisch 2009). In plants with large genomes, LTR retrotransposons often exhibit nested insertional patterns with variation in element orientation (SanMiguel et al. 1996). Such conditions are likely to be conducive to the production of double-stranded RNAs that could trigger the RNAi pathway (Madlung and Comai 2004; Lisch 2009). Future work will be geared toward understanding how and at what point host repression mechanisms are functional in these sunflower populations.

Supplementary Material

Supplementary table S1 is available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

We thank T. Koopman and M. Ty for technical support and we thank Nolan Kane for providing us with microsatellite primer sequences. This work was supported by Kansas State University and National Science Foundation grant (DEB-0742993 to M.C.U.).

Literature Cited

Baack EJ, Whitney KD, Rieseberg LH. 2005. Hybridization and genome size evolution: timing and magnitude of nuclear DNA content increases in *Helianthus* homoploid hybrid species. New Phytol. 167:623–630.

Beaulieu J, Jean M, Belzile F. 2009. The allotetraploid *Arabidopsis thaliana-Arabidopsis lyrata* subsp petraea as an alternative model system for the study of polyploidy in plants. Mol Genet Genomics. 281:421–435.

Bennett MD, Leitch IJ. 1995. Nuclear-DNA amounts in angiosperms. Ann Bot. 76:113–176.

Bino RJ, Lanteri S, Verhoeven HA, Kraak HL. 1993. Flow cytometric determination of nuclear replication stages in seed tissues. Ann Bot. 72:181–187.

Buerkle CA. 2005. Maximum-likelihood estimation of a hybrid index based on molecular markers. Mol Ecol Notes. 5:684–687.

Donovan LA, Rosenthal DM, Sanchez-Velenosi M, Rieseberg LH, Ludwig F. 2010. Are hybrid species more fit than ancestral parent species in the current hybrid species habitats? J Evol Biol. 23:805–816.

Feschotte C, Keswani U, Ranganathan N, Guibotsy ML, Levine D. 2009. Exploring repetitive DNA landscapes using REPCLASS, a tool that automates the classification of transposable elements in eukaryotic genomes. Genome Biol Evol. 1:205–220.

Grandbastien MA. 1998. Activation of plant retrotransposons under stress conditions. Trends Plant Sci. 3:181–187.

Gross BL, et al. 2004. Reconstructing the origin of *Helianthus deserticola*: survival and selection on the desert floor. Am Nat. 164:145–156.

Gross BL, Schwarzbach AE, Rieseberg LH. 2003. Origin(s) of the diploid hybrid species *Helianthus deserticola* (Asteraceae). Am J Bot. 90:1708–1719.

Hazzouri RM, Mohajer A, Dejak SI, Otto SP, Wright SI. 2008. Contrasting patterns of transposable-element insertion polymorphism and nucleotide diversity in autotetraploid and allotetraploid *Arabidopsis* species. Genetics 179:581–592.

Kapitonov VV, Jurka J. 2001. Rolling-circle transposons in eukaryotes. Proc Natl Acad Sci U S A. 98:8714–8719.

Kawakami T, Strakosh SC, Zhen Y, Ungerer MC. 2010. Different scales of *Ty1/copia*-like retrotransposon proliferation in the genomes of three diploid hybrid sunflower species. Heredity 104: 341–350.

Kentner EK, Arnold ML, Wessler SR. 2003. Characterization of high-copy-number retrotransposons from the large genomes of the louisiana iris species and their use as molecular markers. Genetics 164:685–697.

Kumar A, Bennetzen JL. 1999. Plant retrotransposons. Annu Rev Genet. 33:479–532.

Labrador M, Farre M, Utzet F, Fontdevila A. 1999. Interspecific hybridization increases transposition rates of Osvaldo. Mol Biol Evol. 16:931–937.

Lerat E. 2010. Identifying repeats and transposable elements in sequenced genomes: how to find your way through the dense forest of programs. Heredity. 104:520–533.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 25: 1451–1452.

Lisch D. 2009. Epigenetic regulation of transposable elements in plants. Annu Rev Plant Biol. 60:43–66.

Madlung A, Comai L. 2004. The effect of stress on genome regulation and structure. Ann Bot. 94:481–495.

Melayah D, Bonnivard E, Chalhoub B, Audeon C, Grandbastien MA. 2001. The mobility of the tobacco *Tnt1* retrotransposon correlates with its transcriptional activation by fungal factors. Plant J. 28:159–168.

Metcalfe CJ, et al. 2007. Genomic instability within centromeres of interspecific marsupial hybrids. Genetics 177:2507–2517.

Michalak P. 2009. Epigenetic, transposon and small RNA determinants of hybrid dysfunctions. Heredity 102:45–50.

Parisod C, et al. 2009. Rapid structural and epigenetic reorganization near transposable elements in hybrid and allopolyploid genomes in *Spartina*. New Phytol. 184:1003–1015.

Petit M, et al. 2010. Mobilization of retrotransposons in synthetic allotetraploid tobacco. New Phytol. 186:135–147.

Pritham EJ. 2009. Transposable elements and factors influencing their success in eukaryotes. J Hered. 100:648–655.

Rieseberg LH, Baird SJE, Desrochers AM. 1998. Patterns of mating in wild sunflower hybrid zones. Evolution 52:713–726.

Rieseberg LH, et al. 2003. Major ecological transitions in wild sunflowers facilitated by hybridization. Science 301:1211–1216.

Rieseberg LH, Whitton J, Gardner K. 1999. Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. Genetics 152:713–727.

Saha S, Bridges S, Magbanua ZV, Peterson DG. 2008. Empirical comparison of ab initio repeat finding programs. Nucleic Acids Res. 36:2284–2294.

SanMiguel P, et al. 1996. Nested retrotransposons in the intergenic regions of the maize genome. Science 274:765–768.

Schuelke M. 2000. An economic method for the fluorescent labeling of PCR fragments. Nat Biotechnol. 18:233–234.

Schwarzbach AE, Rieseberg LH. 2002. Likely multiple origins of a diploid hybrid sunflower species. Mol Ecol. 11:1703–1715.

Shan X, et al. 2005. Mobilization of the active MITE transposons mPing and Pong in rice by introgression from wild rice (*Zizania latifolia* Griseb.). Mol Biol Evol. 22:976–990.

Slotkin RK, Martienssen R. 2007. Transposable elements and the epigenetic regulation of the genome. Nat Rev Genet. 8:272–285.

Staton SE, Ungerer MC, Moore RC. 2009. The genomic organization of *Ty3/Gypsy*-like retrotransposons in *Helianthus* (Asteraceae) homoploid hybrid species. Am J Bot. 96:1646–1655.

Tenaillon MI, Hollister JD, Gaut BS. 2010. A triptych of the evolution of plant transposable elements. Trends Plant Sci. 15:471–478.

Ungerer MC, Strakosh SC, Stimpson KM. 2009. Proliferation of *Ty3/gypsy*-like retrotransposons in hybrid sunflower taxa inferred from phylogenetic data. BMC Biol. 7:40.

Ungerer MC, Strakosh SC, Zhen Y. 2006. Genome expansion in three hybrid sunflower species is associated with retrotransposon proliferation. Curr Biol. 16:R872–873.

Van den Broeck D, et al. 1998. Transposon display identifies individual transposable elements in high copy number lines. Plant J. 13:121–129.

Verma SP, Quiroz-Ruiz A. 2006. Critical values for six Dixon tests for outliers in normal samples up to sizes 100, and applications in science and engineering. Revista Mexicana De Ciencias Geologicas. 23:133–161.

Vukich M, Giordani T, Natali L, Cavallini A. 2009. *Copia* and *Gypsy* retrotransposons activity in sunflower (*Helianthus annuus* L). BMC Plant Biol. 9:150.

Waugh R, et al. 1997. Genetic distribution of Bare-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). Mol Gen Genet. 253: 687–694.

Waugh O'Neill RJ, O'Neill MJ, Marshall Graves JA. 1998. Undermethylation associated with retroelement activation and chromosomal remodelling in an interspecific mammalian hybrid. Nature 393:68–72.

Welch ME, Rieseberg LH. 2002. Patterns of genetic variation suggest a single, ancient origin for the diploid hybrid species *Helianthus paradoxus*. Evolution 56:2126–2137.

Wessler SR. 1996. Turned on by stress. Plant retrotransposons. Curr Biol. 6:959–961.

Wicker T, et al. 2007. A unified classification system for eukaryotic transposable elements. Nat Rev Genet. 8:973–982.

Zilberman D, Henikoff S. 2004. Silencing of transposons in plant genomes: kick them when they're down. Genome Biol. 5:249.

Associate editor: Brandon Gaut